

Leapfrogging the power grid

The desire to mitigate climate change, and opportunities to empower consumers in the developed and developing worlds, all point towards a need for less-centralized energy generation. It's time to further boost hydrogen research.

You purr your pollution-free hydrogen car into your garage, plug it into the wall, and ponder your electricity meter spinning anticlockwise. The car's fuel cell not only powers your house but, through an intelligent chip, auctions off the surplus electricity to local power brokers. Sooner than you may think, your lab or university might be sending out electricity bills instead of paying them. Technologies, green imperatives and liberalized energy markets are converging to reinvent the century-old electricity grid.

Alongside power plants, oil barons and high-voltage pylons, local generators will cheaply and cleanly pump out power from villages, faculties and backyards, and electricity consumers will also become vendors. Underpinning this shift, computer scientists are creating an intelligent, responsive 'energy web', where local supply and demand would be micromanaged in an electronic energy marketplace.

World energy needs will double by 2050 and we urgently need sources that don't produce carbon dioxide or other pollutants. Decentralized generation puts technological choices in consumers' hands. The technology is available to bring small, local energy sources, or 'micropower', better into the electricity equation, from gas turbines, hydro and wind power and sugar-cane biomass to nanoscale solar cells embedded in the bricks and slates of houses.

For countries like the United States that already have grids, decentralized generation will ease gridlock, radically improve energy efficiency, cut carbon emissions and provide better resilience to failures and terrorist attacks on vulnerable networks.

But for the 2 billion people without electricity, micropower could let them leapfrog the grid. Just as countries that had never seen an expensive copper telephone network jumped straight to mobile phones, so decentralized generation technologies offer the chance for them to leapfrog the grid and prosper. That was the take-home message from a meeting of energy companies, researchers and policy makers in Paris last week (www.wtn.net/new/registration/2004/energy/index.jsp).

Clean and green

The meeting concluded that an array of micropower technologies, from small gas turbines to solar panels, have matured sufficiently to provide electricity to developing countries, even in remote rural areas. Demand for electricity in poor countries is set to rocket. Minimizing the resulting surge in carbon emissions requires them to explore micropower to help them meet demand in cleaner ways and broaden their energy sources to become less dependent on oil.

Progress in industrialized countries will be key not just to convincing poorer countries that this is the way forward, but also to reducing the cost, by first creating a market in the developed world. Distributed generation is already happening, but its share of total electricity production worldwide has remained flat at 7% over the past three years, according to the World Alliance for Decentralized Energy (www.localpower.org). In countries with proactive energy policies, such as Denmark, this share is about 50%.

Efficient gas turbines at the point of use are the most immediately promising form of decentralized generation. Although they largely burn fossil fuels such as natural gas, they are already cleaner and more

efficient than grid electricity. Excluding large hydroelectric facilities, renewables account for only a few percent of total world production, but Denmark and Luxembourg obtain around 17% of their electricity from them. The technological potential is there to expand renewables. Wind is the fastest-growing energy source, accounting for 2% of European needs. The costs of solar energy have plummeted by 90% since the 1970s, bringing it, at around \$4.50 per watt, tantalisingly close to the \$3 per watt of grid electricity — and that's not taking into account the environmental costs and hidden subsidies of the grid.

The ultimate micropower source is the fuel cell, which converts hydrogen to electricity without combustion or moving parts. It overcomes the main drawback of renewable energy sources, that electricity is lost if not immediately used. Excess electricity can be converted to hydrogen and stored to make electricity again. The most immediate cost-effective source of hydrogen is cracking it from natural gas. This would not reduce carbon emissions or the West's dependence on imports, but could provide a stepping stone to the hydrogen economy.

More hydrogen, please

Advocates of hydrogen argue that it has been too glibly dismissed as an energy carrier (see, for example, www.rmi.org/images/other/E-20HydrogenMyths.pdf). A report released earlier this month by the US National Academies — *The Hydrogen Economy: Opportunities, Costs, Barriers, and R&D Needs* — argues that a wholesale shift to a hydrogen economy is still decades away and that major cost obstacles must be overcome, but nevertheless signals a route towards a transition (see www.nap.edu/books/0309091632/html).

As the academy report emphasizes, incremental changes are not enough — making the hydrogen economy a reality will require "a comprehensive, long-range program of innovative, high-risk/high-payoff basic research" in catalysis, nanomaterials, membranes and separation. The report recommends expanding US research to include distributed hydrogen production systems, hydrogen storage and solar energy for hydrogen production.

Distributed generation could go some way to solving the automobile industry's Catch 22 of creating a hydrogen infrastructure: why develop hydrogen cars when there is no distribution network, and why develop a distribution network if there are no hydrogen cars? Fuel cells for generating electricity need to become a common feature of factories, hospitals, fork-lift trucks in Walmart's nearly 6,000 outlets, and fleet vehicles such as city buses, government vehicles and delivery trucks. Pilot initiatives have already been established.

A level playing field is key to more research. Many US states and countries are passing laws that allow customers to sell their electricity, which in the past would have landed you in jail. California's decree that by next year a tenth of all cars sold in the state must not produce emissions is waking up US car-makers. And EU laws require member states to have 12% of all energy renewable by 2010. As fuel cells and other sources of micropower gain market share, market forces will drive down costs and provide research incentives.

Hydrogen as a widely used energy carrier is essential and inevitable. Scientists, technologists, governments and philanthropists should do much more to hasten its arrival. ■





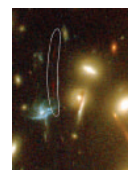
Human cloning
South Koreans
announce embryo
success in Seattle
p664



All change
Rita Colwell steps
down as NSF
director
p665



Space casualty
NASA mission to
search for dark
energy delayed
p667



Distant stars
Call for Hubble
reprieve as it spots
farthest galaxy
p668

Publishers split over response to US trade embargo ruling

Geoff Brumfiel, Washington

Iranians struggling to secure free speech at home are facing a fresh set of restrictions from the US government.

The US Department of the Treasury has ruled that editing or publishing scientific manuscripts from Iran, Libya, Sudan and Cuba violates the trade embargo on these countries. And US publishers and scientific societies are divided over how to respond.

At a meeting in Washington on 9 February, David Mills, the treasury official in charge of implementing the policy, told representatives of 30 publishers that anyone wanting to publish papers from Iran should seek a licence from the treasury department. He also suggested that US scientists collaborating with Iranians could be prosecuted.

The ruling has split US scientific societies. The journals of the Institute of Electrical and Electronics Engineers (IEEE) have stopped accepting manuscripts from researchers in embargoed countries. On the other hand, the American Institute of Physics (AIP), the American Physical Society and the American Association for the Advancement of Science, which publishes



Gagged: Iranian scientists may be unable to publish in US-based journals.

Science, have so far refused to comply. "We feel that we are protected by freedom of speech," says Marc Brodsky, executive director at the AIP.

Questions about interactions with Iran first arose in 2001 when the IEEE tried to rent a conference room at a Tehran meeting, and was told that this would violate the US trade embargo. In ensuing conversations between the organization and the treasury department's Office of Foreign Assets Control, it emerged that publishing could also be restricted. According to a 30 September 2003 letter from the office, editing content from an

author in a restricted country is "prohibited ... unless specifically licensed".

As a result, the IEEE's 100-plus journals began declining manuscripts from researchers in Iran, Cuba, Libya and Sudan. "We felt we needed to operate within the laws of the country we are in," says Michael Lightner of the IEEE. This year, the American Nuclear Society, the American Chemical Society and the American Society for Microbiology followed suit (see *Nature Med.* **10**, 109; 2004). Some of these societies, including the IEEE, are actively pursuing

a licence to publish.

But other groups may seek to overturn the ruling. "The government should not be in the business of restricting this kind of first-amendment activity," says Allan Adler, head of legal and government affairs at the Association of American Publishers, which represents most major for-profit and society publishers in the United States, including the IEEE.

Adler says that the law specifically exempts 'information and informational materials' from trade embargoes. The association is considering several options, including court action and legislation, to overturn the ruling. "We think that this is wrong as a matter of law and a matter of principle," he says.

Nature and other publications of the Nature Publishing Group are still accepting manuscripts from the affected countries. "We see no grounds to absolutely decline to handle papers from these countries," says Philip Campbell, editor of *Nature*. "But we are taking legal advice."

Iranian scientists say that the US policy is crippling their research. "The ruling makes publication by Iranians in journals published in the United States practically impossible," says Fredun Hojabri, an Iranian chemist now living in San Diego, who heads the Sharif University of Technology Association, representing alumni, students and faculty of Iran's premier technical university.

Scientists slam Bush record

Geoff Brumfiel, Washington

Nearly two dozen Nobel laureates and 40 other leading researchers have signed an angry statement accusing the Bush administration of "misrepresenting and suppressing scientific knowledge".

The statement, due to be released this week, accuses the administration of stacking scientific committees, censoring results and dissolving advisory panels. "The systematic, pervasive nature of these practices is unprecedented," says Kurt Gottfried, a particle physicist at Cornell

University and chairman of the Union of Concerned Scientists, which assembled the petition. "The very ethos of science is being challenged," he says.

Signatories include David Baltimore, president of the California Institute of Technology, Harold Varmus, former director of the National Institutes of Health, and Lewis Branscomb, professor of public policy at Harvard University and an adviser to both Republican and Democratic administrations.

■ www.ucsusa.org

Guatemalan forensic work brings award and death threats

Jim Giles, Seattle

"I like Bournemouth," says Fredy Peccerelli of his temporary home, a genteel town on England's south coast. "No one has tried to kill me yet."

In his native Guatemala, death threats were a regular occurrence for Peccerelli, a forensic anthropologist who has spent most of the past decade exhuming bodies from mass graves — the legacy of Guatemala's 35 years of conflict.

Peccerelli got international recognition for his work on 14 February, when he received a human-rights award at the annual conference of the American Association for the Advancement of Science (AAAS) in Seattle.

He is now taking a break to study for a masters degree in anthropology at the University of Bournemouth. But he thinks that the work in Guatemala has only just begun. "In 12 years we have investigated 400 cases and exhumed around 3,000 bodies," he says. "There is enough work for another 25 years."

At least 200,000 people are thought to have been killed in Guatemala during fighting between government forces and left-wing insurgents from the early 1960s to the mid-1990s. Once the conflict receded, Peccerelli, who studied anthropology at Brooklyn College, New York, jumped at the chance to tackle his homeland's problems.

Using funding from charities and the AAAS, the Guatemalan Forensic Anthropology Foundation began responding to requests from rural communities that want to prosecute military officers alleged to have massacred innocent villagers.

The foundation gathers evidence on the attacks and then exhumes the bodies. By determining the cause of death and matching the remains with details of those who have disappeared, Peccerelli's team helps to inform government officials charged with investigating the attacks.

In Guatemala, Peccerelli travels with a bodyguard, as does his family. "I've woken up to find bullet holes in my garage door," he recalls. Other members of his team have been threatened at gunpoint.

Peccerelli is considering further academic studies, although he continues to raise money to keep the investigations going. "I need some stability in my life," says Peccerelli. "I might have to change my work for my own sanity." ■



Healthy culture: Koreans support stem-cell research, say Shin Yong Moon (left) and Woo Suk Hwang.

Cloning success marks Asian nations as scientific tigers

Helen Pearson, Seattle

The successful cloning of human embryos by a South Korean team has alerted Western researchers to the pace of scientific and technological progress in East Asia, biologists say.

A team led by Woo Suk Hwang and Shin Yong Moon of Seoul National University last week revealed that they had grown the first embryonic stem-cell line derived from a cloned human embryo (W. S. Hwang *et al.* *Science* 10.1126/science.1094515).

Although the lead researchers are already renowned in their own country for their cloning and stem-cell work, the result has highlighted research in nations such as South Korea and Singapore. "Asian workers are likely to be major players in this field," predicts stem-cell pioneer Roger Pedersen of the University of Cambridge, UK.

Hwang and Moon attribute their success to a supportive cultural environment, well-funded laboratories, and legislation that permits the cloning of human embryos for research.

Hwang adds that the rarity of organ donation in South Korea lends impetus to research that might create alternative sources of transplant tissue. The work ethic in the labs also helps, he says: "There is no Saturday, no Sunday and no holidays," he told the annual meeting of the American Association for the Advancement of Science (AAAS) in Seattle on 16 February.

Critical to the researchers' success was their collection of 242 eggs from 16 female

volunteers, with which they fine-tuned the tricky cloning method. They injected the nucleus from one of the cumulus cells that nestles around a woman's eggs into an egg that had been stripped of its genetic material. About a quarter of the embryos grew into blastocysts, and the team extracted and grew embryonic stem cells — which can generate many different tissues — from one of these.

While lauding the advance, some at the AAAS meeting said that the pace of human therapeutic cloning research in Asia and elsewhere threatens to leave US scientists stranded, as they cannot get federal funds to derive stem cells from human embryos.

But some US biologists say that they can contribute to the field by collaborating with researchers in South Korea, Britain and other countries where the work is supported. "We should behave in a complementary manner," says Gerald Schatten, who studies primate cloning at the University of Pittsburgh. "We don't have to do everything in every country."

The cloning breakthrough has revived public debate in the United States on whether such research should be taking place. There is no US legislation on human cloning because of a stalemate between those who want to outlaw all forms of human cloning and those who want to ban reproductive cloning but allow therapeutic cloning. Because the question is linked to abortion, it could well emerge as an issue in this year's presidential election campaign, specialists at the meeting said. ■

Colwell calls time on 'wonderful run' at NSF

Geoff Brumfiel, Washington

Rita Colwell will step down on 21 February as director of the National Science Foundation (NSF), the agency that supports most non-biomedical research in the United States, after five-and-a-half years in charge.

Colwell, the first woman and only the second biologist to lead the NSF, announced her departure at the 11 February hearing of the House Committee on Science. "It's been a wonderful run," she said afterwards. Arden Bement, a metallurgist who heads the National Institute of Standards and Technology, will run the agency until the White House chooses a permanent successor.

Colwell will now head up Canon Life Sciences, a new offshoot of the electronics company. She is resuming her professorship at the University of Maryland, College Park, and taking one up at Johns Hopkins University in Baltimore, where she will help develop a centre for the study of infectious diseases.

During her tenure, the NSF's budget grew by two-thirds, from US\$3.4 billion in 1998 to \$5.6 billion this year. The average annual NSF grant rose from \$90,000 to \$142,000, she points out, and graduate student stipends almost doubled, to \$30,000. "She's certainly been effective," says one Senate staffer.

More controversial were the millions of dollars spent on large projects, including \$300 million for a gravity-wave detector and



Rita Colwell, the first woman to run the NSF, presided over a budget rise of two-thirds in six years.

\$219 million for a network of geological observatories.

The management of these projects was criticized in the Congress (see *Nature* 418, 573; 2002) and Colwell's insistence that all of them were progressing satisfactorily irritated some. "She drove them nuts from time to time," says Joel Widder, a lobbyist and former senior NSF official who has worked for a congressional committee overseeing the agency. "She would continually insist that everything was on time and on budget."

And Colwell, a marine biologist, had

mixed results in raising money for her first love, environmental and ecological research. In 2000, she launched an interdisciplinary initiative on biocomplexity, which to date has received \$334 million. But she was unable to implement a much larger move by the agency into environmental science proposed in 1998 (see *Nature* 395, 4; 1998).

Some were surprised at an outsider being appointed as the agency's interim leader. But Bement is seen as a safe pair of hands — he knows the NSF well, having served on its governing board for six years from 1989. ■

Medical research wins simpler grants and extra cash

Laura Nelson, London

UK researchers are to get more cash and a simpler grant system, after a budget overhaul at the Medical Research Council (MRC).

The changes, which many researchers have been demanding for years, were announced on 13 February by Colin Blakemore, the MRC's chief executive. Blakemore, a neurobiologist, arrived to run the council from the University of Oxford last October.

In the financial year starting this April, the MRC will have £144 million (US\$270 million) for competitive grant proposals, up from £124 million this year. And instead of having to apply in large, cooperative groups, investigators can go it alone.

Some researchers, and the House of Commons' science and technology committee, had criticized the old scheme,

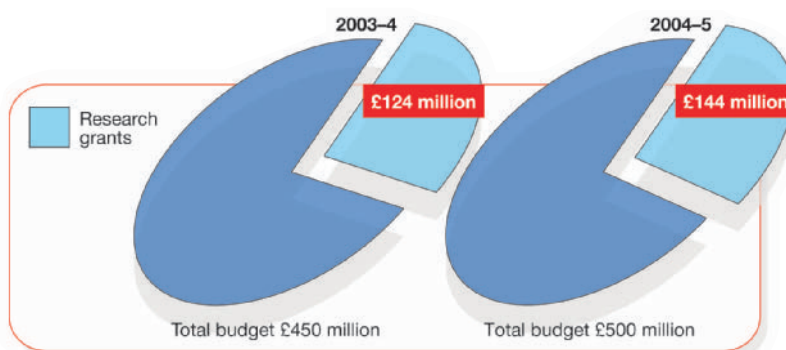
developed under Blakemore's predecessor George Radda (see *Nature* 422, 461; 2003).

The new policy is a huge improvement, says David Price, a physiologist at the University of Edinburgh. The cooperative arrangement was "incredibly bureaucratic", he says. "Now we can think about the science we have in mind, rather than looking through the MRC schemes and modifying our science to suit them." Radda declined to comment on the change.

The MRC, the second-largest of the seven UK research councils, will provide fewer types of grants, but will offer more flexibility in their size and duration. Grants will vary from two to five years, and Blakemore hopes there will be a shortening of the application process from 26 to between 20 and 22 weeks.

Blakemore says that scientists and universities were consulted extensively in drawing up the reforms. "It was clear that simpler and more flexible schemes were needed," he says.

Large MRC projects, such as a planned genetic database called Biobank, won't be squeezed by the change, he adds. But some scientists say that they ought to be. "I am concerned about the way funding is skewed towards big projects," says Jim Cohen of the MRC Centre for Developmental Neurobiology at King's College, London. ■



Misconduct row fuels calls for reform

Carina Dennis, Sydney

A fierce row over misconduct allegations has prompted Australian researchers to call for an office of research integrity to be set up.

The issue came to a head on 10 February, when Senator Kim Carr, research spokesman for the opposition Labor Party, released to parliament an unpublished report of an inquiry into allegations made about an immunologist at the University of New South Wales.

The case centres on Bruce Hall, a transplant immunologist working on graft tolerance. In 2001, Hall was accused by four complainants in his laboratory of fabricating and falsifying experimental results in an abstract and paper, providing false data in a grant application, misattributing author-

ship credit and workplace bullying. The experiments in question involved the use of cells from the immune system — called CD4⁺ T cells — to transfer graft tolerance to rats with transplanted hearts. Hall emphatically denies the allegations.

In June 2002, the university commissioned an external panel chaired by Gerard Brennan, Australia's former chief justice, to investigate. The panel's report, delivered to the university in January 2003 but not published until Carr released it, said that Hall had engaged in scientific misconduct.

Hall himself hotly contests the findings of the Brennan report. "They got the science wrong," he says, adding that panel members did not understand his field of research. Before the report's release in parliament,

Hall had taken court action to try to block its publication. And last December, Rory Hume, vice-chancellor of the university, ruled that Hall was not guilty of scientific misconduct and decided to censure him formally for lesser charges of academic misconduct.

But Carr has criticized the university for not releasing the report earlier. "It's a matter of public interest," says Carr. "I don't believe the university has responded appropriately." Several members of the university's governing council, as well as some faculty members, have also taken issue with the handling of the matter.

For many researchers, the greatest concern is not so much the individual case but the flaws it exposes in the way Australian universities handle misconduct cases. Several have called for an office of research integrity, similar to the one in the US health department. "The current process is inadequate," says Warwick Anderson, head of Monash University's school of biomedical sciences. "We've now been faced with a major case of a reputable investigator and a highly respected university — it has been deeply problematic and hasn't served anyone's interests." Hume agrees: "Australia could do well to have an office of scientific integrity," he says.

The National Health and Medical Research Council, Australia's main funding agency for biomedical research, terminated Hall's grant last October, saying that he had provided misleading information in his grant application. In June 2002, the agency initiated a review of Australia's guidelines on all aspects of the conduct of research, which is being conducted jointly with the Australian Research Council and the universities. ■



Question time: scientific misconduct has been brought to the fore in Australia's parliament.

Berlin biologists outraged by imminent pay cut

Anna Wellmann, Munich

Biologists at one of Berlin's leading biomedical research centres have been told to prepare for a 12% pay cut. Unions representing 700 staff at the Max Delbrück Center for Molecular Medicine (MDC) say that they will fight the cut in the courts.

Berlin's government and the federal research ministry, which finance the centre jointly, say that the city's fiscal crisis leaves them no choice but to make the cut. Facing a yawning budget deficit, Berlin decided last August to cut the salaries of its public employees in exchange for a shorter working week and guaranteed jobs until 2009.

The capital's three universities and some other research institutes were exempted, but the MDC was not. The federal research ministry approved the

pay cuts at the beginning of this month.

Staff are not amused. They say that, as researchers, they are unlikely to benefit from cuts in their hours. And the guarantee of employment doesn't help, as they are nearly all on five-year, fixed-term contracts.

"I don't know how my family could cope with 12% less per month," says Thomas Müller, a postdoctoral fellow studying developmental biology at the centre. Udo Heinemann, a crystallographer there, adds that lower salaries will make it hard for the centre to attract staff.

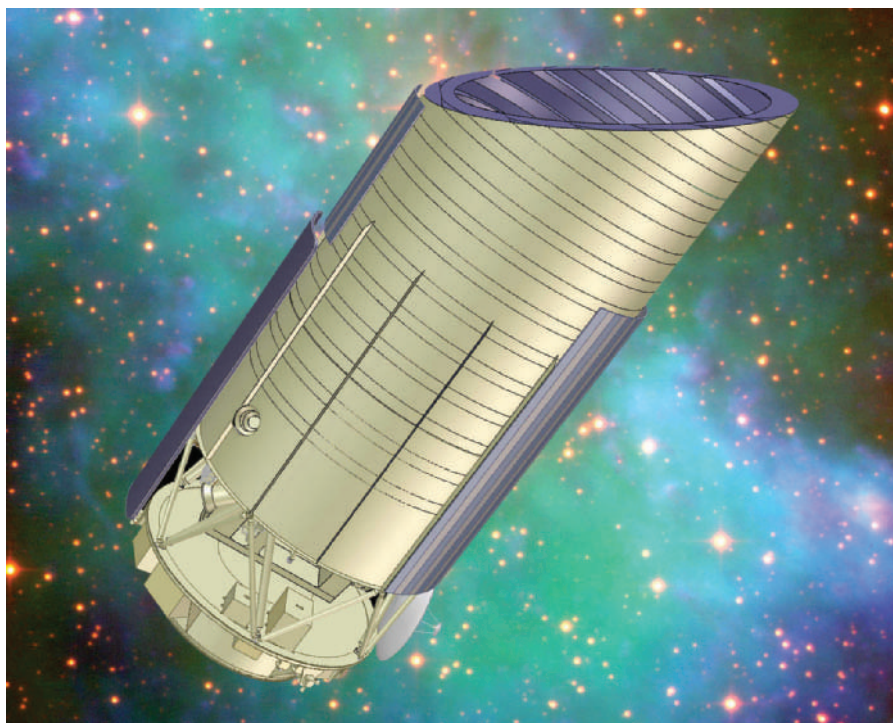
As a compromise, the research ministry has suggested giving MDC employees the same terms as other public workers in eastern Germany. This would result in a 7.5% pay cut, but no reduction in the working week. Marion Bimmler, who chairs



Researchers at Berlin's Max Delbrück Center plan to contest a sharp wage cut in the courts.

the staff council of the MDC, compares this to "choosing between plague and cholera".

Bimmler says the unions are advised that the pay cut is illegal. But court action could last for more than a year, and the new terms are taking effect immediately. ■



Under threat? Probes designed to investigate deep space, such as SNAP, face unexpected delays.

NASA casts a shadow over bid to illuminate dark energy

Geoff Brumfiel, Washington

NASA is delaying a mission to explore the fundamental make-up of the Universe, riling astronomers and physicists who consider it a top priority.

About three-quarters of all the mass and energy in the Universe may be dark energy. In a ground-breaking collaboration, NASA and the US Department of Energy (DOE) were set to explore this poorly understood entity on the Joint Dark Energy Mission, which supporters hoped would be launched into space by 2014 (see *Nature* 425, 887; 2003).

But the collaboration is now on hold, as NASA re-tools for possible manned flights to the Moon and Mars. Two other missions in NASA's 'Beyond Einstein' astrophysics programme — one designed to probe microwaves left over from the Big Bang and the other to search for black holes — also failed to show up in NASA's five-year budget outlook, released on 2 February. Physicists had expected the missions to be in the plan if they were to be completed on the schedule that the space agency allocated to them last year.

The dark energy mission was listed as a top priority in a National Academy of Sciences study of interdisciplinary research in astronomy and physics, published in April 2002. "It's a serious casualty," says Roger Blandford, who heads the Kavli Institute for Particle Astrophysics and Cosmology

at Stanford University in California.

The DOE remains committed to the project — at least for now. Saul Perlmutter, a cosmologist at Lawrence Berkeley National Laboratory in California, is the primary investigator for one proposed version of the mission — the supernova/acceleration probe, or SNAP. He says that he is disappointed by NASA's decision but adds that the research can progress without the space agency's involvement. Robin Staffin, director of the DOE's high-energy physics division, says the department will spend between \$7 million and \$8 million next year developing detectors for the mission.

But researchers acknowledge that it will ultimately require NASA's support to carry out a mission that is expected to cost about \$900 million.

NASA programme manager Paul Hertz downplayed the decision to omit it from the agency's five-year budget plan. "Nothing was cancelled," he says. "But the timeline has been stretched out." He says that the Office of Space Science, which had overseen the project, will continue to push for the mission to go ahead.

But Rocky Kolb, who was part of an independent NASA advisory board that originally recommended the mission, says it may prove a tough sell. "The immediate problem is that no one knows how to use dark energy to get to Mars," he quips. ■

Thumbs up for fresh formula to gauge university funding

Jim Giles, London

What's the best way to divide £8 billion (US\$15 billion) in annual funding between Britain's 100-plus universities? A fresh approach to the challenge was revealed last week, and looks to have the qualified backing of academics and administrators.

The plan would revamp the Research Assessment Exercise, one of the world's most extensive attempts to measure the performance of university departments. The UK government uses the findings to allocate block grants that subsidize research and teaching at each university.

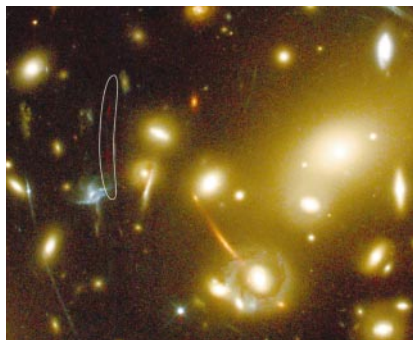
But the last assessment, in 2001, drew sharp criticism. So the plan would scrap its most contentious component — a seven-point ranking scale for every university department that tended to reward small variations in performance with big differences in funding.

In its place, the funding councils will draw up a more detailed profile of each department's output. Each piece of work submitted to the exercise, such as a published book or paper, will be awarded between zero and four stars, with the top two rankings corresponding to work of international significance. The profile will show what percentage of submissions achieved each ranking, and how many researchers entered work.

The councils say these profiles will provide a more subtle snapshot of performance, allowing money to be distributed more fairly. They also hope the profiles will help identify pockets of excellence in average departments — something that was lost when scores were averaged to create the old rankings. "The true scale and strength of the best work will be more visible," says Howard Newby, chief executive of the Higher Education Funding Council for England.

University administrators have praised the councils for ditching some of the proposals made last May in an outside review of the assessment exercise (see *Nature* 423, 574; 2003). A proposal to allow less research-intensive institutes to go through a 'light-touch' assessment has not been taken up, nor has the idea of running interim evaluations between main assessments, which take place roughly once every six years.

But one gripe hasn't been resolved — the funding councils have denied requests from universities to publish the formulae that link their profiles to the amount of money they get. ■



Far, far away: a galaxy 13 billion light years from Earth lies in the region ringed on the left.

Astronomers argue for Hubble reprieve in light of distant galaxy

Seattle Researchers protesting against the retirement of the Hubble Space Telescope are welcoming some fortuitous timing: Hubble has spotted the farthest known object in the Universe.

The object is a galaxy thought to be about 13 billion light years away. The discovery was announced on 15 February at the annual conference of the American Association for the Advancement of Science in Seattle. Hubble is critical for studying

such distant objects, astronomers at the meeting said. They urged NASA to reconsider its decision to scrap the telescope around 2007.

The galaxy was discovered by Jean-Paul Kneib and his colleagues at the California Institute of Technology, Pasadena, who will publish their report in a forthcoming issue of *The Astrophysical Journal*.

Pacific nations splash out on humpback whale survey

San Diego A leviathan survey of humpback whales began this month in the Pacific Ocean, with the goal of determining the animals' population structure.

The survey, funded by the US, Canadian and Mexican governments to the tune of about US\$2 million, will chart humpbacks for three years in their habitats from Mexico to Japan. Researchers plan to conduct genetic studies and examine how the whales are harmed by ships, fishing nets or toxins.

Called the Structure of Populations, Levels of Abundance and Status of Humpbacks (SPLASH) study, the project will help to determine whether the humpback's status as an endangered species under the US Endangered Species Act should be changed.

By the late 1960s, hunting had reduced the humpback population down to about

1,500 animals, scientists say. There are now estimated to be about 10,000 individuals worldwide.

Beagle inquiry gets off the launchpad

London The British government and the European Space Agency (ESA) are to launch an inquiry to determine what went wrong with Europe's first attempt to land on Mars.

The martian probe Beagle 2, supposed to land on Christmas Day 2003, is known to have detached successfully from its mother ship, Mars Express, but failed to send its signature signal, a nine-note tune composed by rock band Blur. It is not known whether the lander made it to the surface or not.

David Southwood, ESA's head of science, said that although he did not want to preempt the inquiry's findings, the Beagle team was "terribly distracted at critical parts of the mission, and delayed by running around trying to pull the support together".

UK science and innovation minister Lord Sainsbury says that the inquiry, which will be headed by ESA inspector general René Bonnefoy, "will allow the experience gained from Beagle 2 to be used for the benefit of future European planetary exploration missions". The report is due by the end of next month.

European patent win aims to protect public research

London The research charity Cancer Research UK last week won a European patent on the *BRCA2* breast-cancer gene, a move intended to guarantee the gene's free availability for research and tests in Europe.

The patent, based on research by Mike Stratton and his colleagues at Britain's Institute of Cancer Research, covers methods for sequencing the gene and testing for damaged and inactive variants, and as such broadly covers much *BRCA2* research. Women with a damaged *BRCA2* gene have higher risks of developing breast and ovarian cancer.

Cancer Research UK will license the patent free to public labs, thereby restricting the activities of private companies such as the Utah-based biotechnology company Myriad Genetics, which holds other patents on the gene and has required that all diagnostic tests be done by the company (see *Nature* 413, 95; 2001).

Élite institute to be spread around Italy

Genoa A trio of Nobel laureates will help shape the Italian Institute of Technology, the elite organization announced by the



Prize catch: Paul Greengard and Rita Levi-Montalcini will help form an Italian institute.

government last year. But the facility's planned form has already changed.

Paul Greengard, Harold Varmus and Rita Levi-Montalcini — all winners of the Nobel Prize for Physiology or Medicine — will sit on the institute's advisory council, most of whose members are non-Italian scientists.

The government has ditched its original idea that the institute should be autonomous, after pressure from Italy's notoriously underfunded universities and research organizations (see *Nature* 426, 111; 2003). They argued that the money would be better spent improving existing bodies. Now, the institute will be “integrated in the

Italian research network”, Italian research minister Letizia Moratti said on Monday.

The government has earmarked €1 billion (US\$1.3 billion) for the scheme over ten years, of which €50 million is available in 2004. The institute will be based in Genoa.

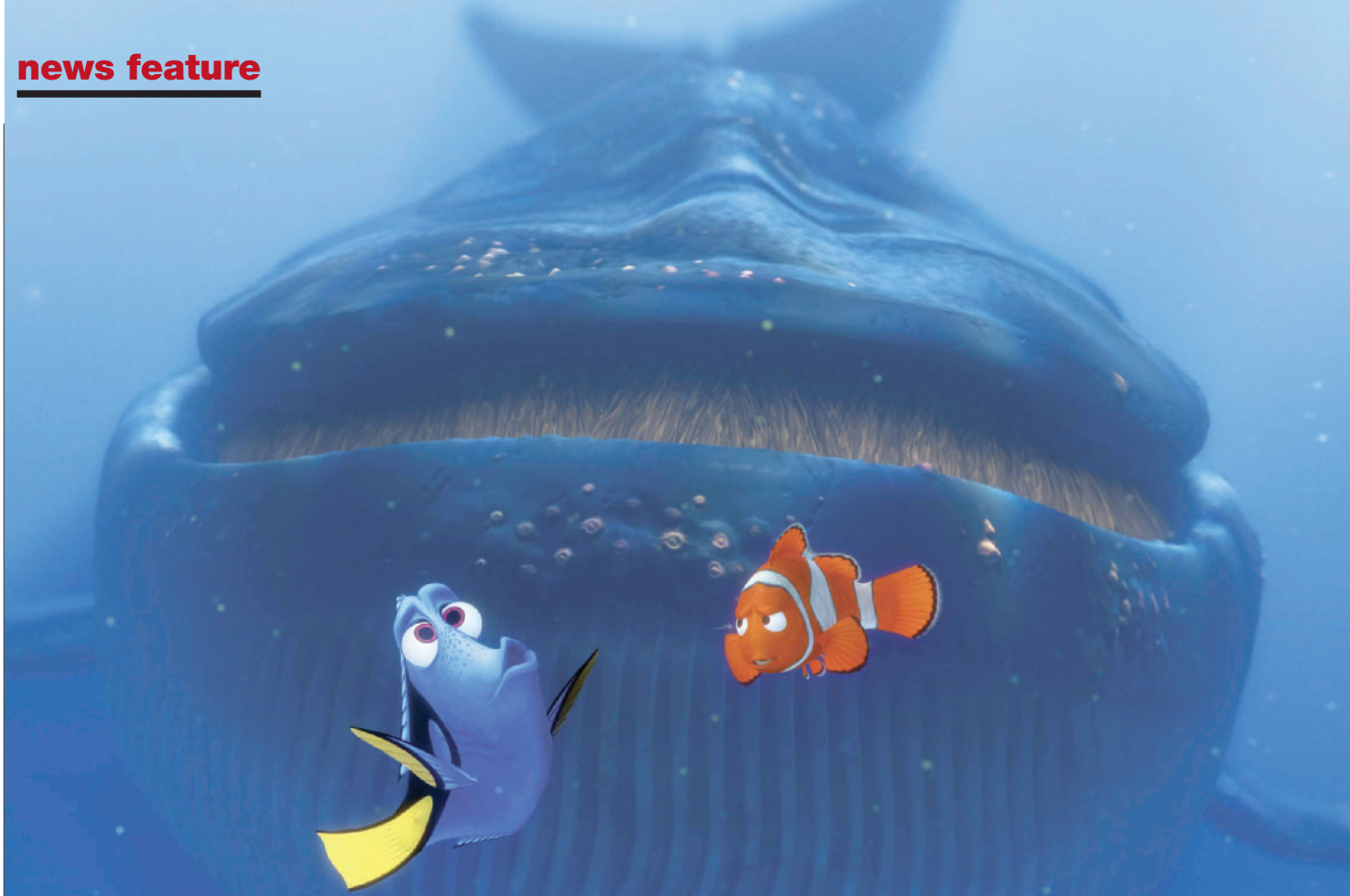
Climate change poses grave threat to corals

Seattle Only sweeping measures can stop global warming doing irreparable harm to coral reefs, marine experts have warned.

Almost 15% of the world's reefs are damaged beyond repair, and another 30% may be lost over the next 30 years, says a report published by the non-profit Pew Center on Global Climate Change. Reefs harbour about 25% of marine species and generate some US\$30 billion each year through fishing and tourism.

The report, released this week at the American Association for the Advancement of Science meeting in Seattle, urges governments to curb greenhouse-gas emissions and to cut the impact of fishing and pollution.

Some countries are already conserving their reefs. In Australia, a law banning fishing on 30% of the Great Barrier Reef is expected to come into effect this year.



The fabulous fish guy

Last year's movie smash *Finding Nemo* impressed many marine biologists with its scientific accuracy. Alison Abbott meets the young expert in fish biomechanics who helped to breathe life into the film's stars.

Adam Summers doesn't own a television. And he had never seen an animated movie until a chance encounter propelled him into the studios of Pixar, as scientific adviser on the company's film, *Finding Nemo*.

The film tells a stirring tale of the efforts of Marlin, a widowed barrier-reef clown fish, to rescue his only son Nemo who has been 'abducted', and taken to live in a dental surgery's aquarium in Sydney. It was the highest-grossing film of 2003, and is in contention for four Oscars at the Annual Academy Awards on 29 February.

In early 2000, Summers was beginning a postdoc in fish biomechanics at the University of California, Berkeley. He rented an apartment owned by a woman who was an art teacher at Pixar. At that time, the studio had just approved a script starring a fish—and the director and animators urgently wanted to know more about their subject matter.

Summers' landlady acted as a go-between, and invited him to give a lecture.

Summers didn't know quite what to expect when he first stepped in front of director and scriptwriter Andrew Stanton and his team in the luxurious screening rooms of the Pixar studios in Emeryville, California. But the film-makers devoured his words like sharks in a feeding frenzy. Summers told them about fish locomotion, behaviour, physiology and coloration. "It was the most engaged class I've ever taught," he says. "I could only get out two or three sentences before a hand would shoot up—and it was graduate-level stuff."

Tales from the deep

Summers also talked about fish oddities, such as the deep-sea anglerfish that live in such darkness that finding a mate is difficult. The tiny males swim up trails of pheromones left by the large females, and then latch on to them so tightly that they

eventually grow into each other. The male becomes little more than a parasitic testicle, able to fertilize eggs whenever the female is ready to lay them. "They loved these stories," says Summers, who is now an assistant professor at the University of California, Irvine. "My one-hour lecture stretched to two and a half hours."

He thought that would be it, but two days later, his landlady knocked on his door again: the studio wanted more. And so Summers found himself giving a three-year course in ichthyology. He organized some 20 lectures on subjects ranging from swimming mechanics to the social behaviour of fish. Some of the topics he suggested himself; others were requested by Pixar employees. All of the talks were recorded so that newcomers could catch up, but Summers was still asked to repeat some lectures—such as his popular discourse on fish locomotion—two or three times.

"In every movie, you need as much research as possible: for every fact you use, you have ten more you need to know about," says Stanton. "We had to traverse a whole ocean in the movie, so we needed a lot of knowledge."

Summers even organized a few makeshift 'labs'. Light quality is immensely important for animators, particularly the 'shaders' who have to worry about how light is reflected from surfaces. So they wanted to know, in huge detail, how fish scales reflect light. Summers brought in a selection of different fish and some microscopes, set up trestle tables and launched into a practical dissecting class



Making a splash: Marlin the clown fish's fantastic voyage was made more realistic thanks to Adam Summers (left) — although the story did bend some rules for sharks and whales.

research to extremes. It was a long swim to Sydney, so Marlin was helped on the last leg of his journey by a passing whale, who swept him into his vast mouth; later the fish was expelled through the whale's blow hole into the city's harbour. Cooper told Summers that she needed first-hand experience of the texture of the inside of a whale.

A whale of a time

Berkeley's Museum of Vertebrate Zoology, where Summers was working, is always alerted when creatures wash up from the sea. So he was able to take Cooper to look inside a newly beached dead grey whale, which had appeared on the shore near Marin, north of San Francisco Bay. Venturing her camera-bearing arm through the whale's blowhole, its mouth, and even bits of its rotting flesh, Cooper took scores of digital shots. The experience gave her a perfect feel for the reflective properties of the inside of a whale. "But she ended up quite smelly," Summers recalls.

Marine biologists appreciate the film for its lack of obvious scientific inaccuracies. "Most of my colleagues have seen it and loved it — there's some front-line fish biology in there," comments Bob Cashner, a former president of the American Society of Ichthyologists and Herpetologists, and a vice-chancellor at the University of New Orleans.

"I'm just amazed at how rigorous these people were," says Summers. One detail involved a female anglerfish whose fluorescent 'lure' — a protuberance extending from its dorsal fin used to attract prey — was used by Marlin to help him find the goggles of

Nemo's abductor in the murky depths. As described in Summers' initial lecture, she has a parasitic male clamped onto her body, just above her anal fin.

After one talk, Summers recalls, when *Nemo* was well along the production road, the director asked him and his guest lecturer, Mike Graham of the Moss Landing Marine Laboratories on Monterey Bay, California, if there was one thing that the film might get wrong that would really disturb them.

Quick as a flash, Graham said the most intolerable outrage would be to see kelp — a type of seaweed that grows only in cold waters — depicted in a coral reef. There was an uncomfortable shuffling in the audience. Then a voice from the back called out: "Better not go see the movie, then." But if you check out your video or DVD, you'll see that there is no kelp. After Graham raised his objection, every frond was carefully removed from each scene, at considerable cost.

Artistic licence

Sometimes, though, it was not possible for the film-makers to be true to the science and keep the story moving. Take Marlin's expulsion into Sydney Harbour. In fact, there is no connection between a whale's mouth and its blowhole. But Nemo's father had to get out into the harbour somehow, and whales can't spit. So in order to drive the plot, Stanton decided to defy nature.

Another compromise: the film's male sharks had no claspers — the fin extensions used to direct sperm into females during mating. It was not a matter of prudery — Summers' objections were silenced when the film-makers showed him that the sharks needed to be depicted much shorter than reality so that they would remain recognizable to viewers as they moved and span. "Because of the shortening, the claspers would have looked like umbrellas trawling around after them — so of course they had to go," says Summers.

It was a similar story with the hammerhead shark's nostrils, which should have been positioned at the far ends of the hammer. This would have left the character unable to make the facial expressions required for his role. Those fish that didn't have to act — the extras, if you like — had more realistic faces. "The 'lies' were always conscious, always a trade-off with the story," says Summers.

Indeed, so concerned were the film-makers about accuracy that Summers remained in hot demand throughout the production process. And although he won't be at the Oscar ceremony, he has been rewarded with a morsel of immortality. Right at the end of the film's long list of credits, below the voices, the technical directors and the caterers, comes: "Adam Summers — Fabulous Fish Guy." ■

Allison Abbott is Nature's senior European correspondent.

► www.pixar.com/featurefilms/nemo

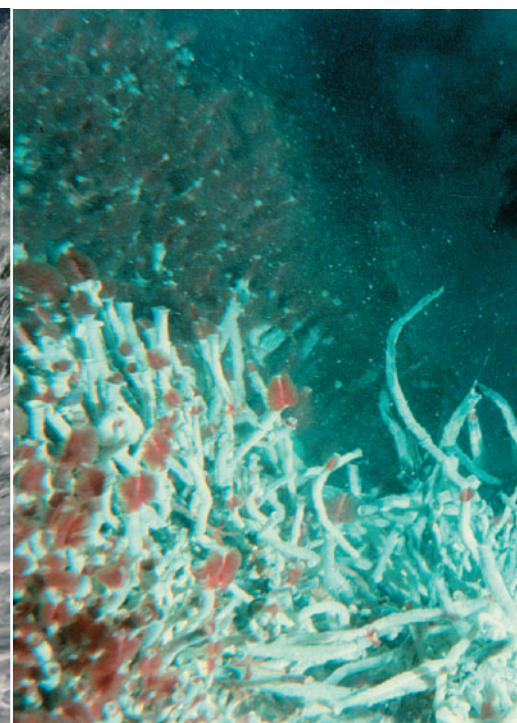
► ecoevo.bio.uci.edu/Faculty/Summers/Summers.html

to explain how the optical properties of fish scales can give rise to 'structural' colours. In another lab, his pupils dissected fish heads to understand the limits of jaw movement.

Summers enlisted help for some of the lectures. "I didn't feel comfortable lecturing on some topics outside my personal expertise to such an inquisitive audience," he says. He invited experts to teach on whales, on the mechanics of waves, and on jellyfish movement and taxonomy.

Stanton was a constant presence, and attendance grew with each lecture to around two dozen people — animators, shaders, colourists, programmers, producers, writers and character developers all joined their director at various times. "We lucked out with Adam," says Stanton. "He was so enthusiastic as well as knowledgeable."

Head shader Robin Cooper took her



Born in a watery commune

If you go back far enough, humans, frogs, bacteria and slime moulds share a common ancestor. But scientists can't agree what it was like, or even whether it was a single creature. John Whitfield reviews the evidence.

"Probably all of the organic beings which have ever lived on this Earth have descended from some one primordial form," Darwin wrote in his *Origin of Species*, published in 1859. Darwin had no way to peer that far back in time. But genome sequencing has given researchers hope that they can finally learn something about the ancestor of all life. In 1999, they even gave it a name, LUCA, for the last universal common ancestor¹.

Yet despite the wealth of genomic data, LUCA has proven elusive. In theory, remnants of the organism from which all life evolved should be scattered around modern genomes. But so far, efforts to reconstruct LUCA's genes by building family trees from modern sequences have ended in frustration. Basic questions about LUCA's nature remain unanswered. Did it live in a hot-water environment, such as a hydrothermal vent at the bottom of the ocean, or in cooler conditions at the ocean surface? Was LUCA simple, like a bacterium, or more complex?

But there are now hints that some of these questions may be answerable. While some groups are zeroing in on LUCA's preferred temperature, others are targeting its genetic blueprint. From all this work, one of the more surprising theories to emerge may also help to explain why LUCA has been

so hard to find. Perhaps it wasn't a single organism at all. Instead, most researchers now believe we should think of LUCA as a pool of genes shared among a host of primitive organisms.

"The naive picture that a group of organisms got all their genes from a simple last common ancestor is breaking down," says microbiologist Gary Olsen of the University of Illinois at Urbana-Champaign. In its place, the image of a sophisticated, global community is emerging, he says. "In the past two years, it feels like it's fallen together into a coherent picture." Rather than a last common ancestor, LUCA may have been a last common community.

Family fortunes

The search for LUCA began with the study of ribosomal RNA, an essential part of the cell's protein-making machinery. The genes for these molecules are highly conserved, meaning the sequences of their four bases (cytosine, guanine, adenine and thymine, which becomes uracil in RNA) are almost identical in all forms of life. Because these sequences have changed so little over time, they are ideal for building family trees, or phylogenies, of evolutionary splits that occurred billions of years ago.

The technique is powerful, but it has not

yet given a definitive answer to the question of temperature. Some believe LUCA was 'thermophilic' or heat-loving. Not only was Earth probably warmer 3.5 billion years ago², but the discovery in the 1970s of a new domain of life, the archaea, also suggested a hot-water origin³. These single-celled organisms share characteristics with the other two domains, the bacteria and the eukaryotes — more complex creatures such as fungi, plants and animals. In many phylogenies, the modern groups most closely related to the common ancestor of archaea and bacteria are hyperthermophiles that live in places where temperatures exceed 80 °C — for example the hot springs of Yellowstone National Park in Wyoming and hydrothermal vents on the ocean floor.

"No matter where you root the tree, all of the groups closest to the bottom are hyperthermophiles," says microbiologist Michael Adams of the University of Georgia, Athens.

Others, however, think LUCA might have preferred the cooler surface waters of the ocean. They distrust phylogenies that point to hot origins because they reach so far back. Mutations build up and overwrite one another over such a span of time, erasing the phylogenetic signal. Also, phylogenetic analysis has some statistical quirks. For example, the gene sequences that evolve



JEFF VANUGA/CORBIS

Blowing hot and cold: genetic analysis has failed to resolve the question of whether the last common ancestor of bacteria, eukaryotes and archaeans lived in cool, shallow waters (left), or in a hot environment such as hydrothermal vents (centre) or hot springs like those in Yellowstone National Park (right).

most quickly tend to come together on trees, even if they are only distantly related. A recent study based on slowly evolving sequences placed a non-thermophilic group at the base of the bacterial family tree⁴.

Looking at LUCA's ribosomal RNA could also reveal something about the temperature of its habitat. Organisms in hot environments have ribosomal RNA that is rich in pairs of the bases guanine (G) and cytosine (C). These G–C pairs are more tightly bound than the alternative, A–U, and are therefore more stable at high temperatures.

But so far the results have been split. One reconstruction of a putative ribosomal RNA sequence for LUCA suggests that it was a cool-water creature⁵. But this analysis omitted some thermophilic groups, says David Saul of the University of Auckland. He believes that there is a trend towards lower G–C ratios as one moves up the tree of life, showing that LUCA was a thermophile. So whether LUCA played in the surf or basked in the hot springs is still an open question.

It had been hoped that ribosomal RNA would also reveal which of the three domains of life — bacteria, eukaryotes or archaeans — LUCA was most like. It was not so simple. Building phylogenies relies on the principle that a bigger difference in sequences between two species means a more remote common ancestor. But the number of possible trees rises exponentially with each species added to the analysis. For deep phylogenies, which reach into the distant past, ribosomal RNA sequences from dozens of species are required for comparison, and the number of possible trees is astronomical. Researchers

have devised mathematical techniques to find the most likely tree, but it is often difficult to choose between the many possibilities with any confidence.

Comparing several genes can make the choice easier, and over the past decade, high-throughput genome analysis has begun to make this feasible. More than a hundred organisms have now been fully sequenced, leading to a great deal of optimism that the roots of life might be traceable. After all, the genetic code and much of biochemistry is universal, so perhaps LUCA can be revealed by finding a set of genes for basic biological functions that is present in all organisms.

Universal genes

Oddly, this extensive comparison of genome sequences from widely divergent modern organisms has identified only about 60 genes that appear to be universal, and therefore probably date back to LUCA. That's nowhere near enough to sustain an organism, says Eugene Koonin, an evolutionary genomics researcher at the National Center for Biotechnology Information in Bethesda, Maryland. The majority of these genes are involved in translation⁶, the process of converting the sequence of bases in DNA into the sequence of amino acids in protein. "On these genes alone, LUCA would go nowhere," Koonin says. "There is nothing for a cell membrane, or for energy metabolism, or any synthetic capabilities. There should have been several times more genes."

According to some evolutionary biologists, the implications for LUCA are strange indeed. If a single LUCA laid the foundations

for the modern diversity in membranes, metabolism and so on, it must have had several different versions of many important genes, in addition to the universal 60. Later lineages would each have pruned all but one from this set, giving rise to the current diversity in basic biochemical pathways. The idea that organisms become more complex rather than less as you get closer to the root of the tree of life is impossible to swallow, says Saul. A single LUCA "would have to have had the most bizarre biochemistry imaginable".

At the same time, evidence was mounting that early life forms were particularly promiscuous in sharing their genes around, in a process called horizontal transfer. Among the genes that should be highly conserved — and therefore good for phylogenies — are those involved in handling genetic information, such as DNA polymerase, which copies DNA, and topoisomerase, which controls the structure of DNA. But the surprise result from this work was that the patterns of ancestry vary depending on which gene you look at⁷. In other words, the phylogenies revealed only the ancestry of the genes themselves, not the relatedness of the species that housed them. This showed genes had hopped between lineages.

For those trying to trace the tree of life back to its roots, the evidence that horizontal transfer had scrambled phylogenies was a serious blow. The impact on evolutionary biology was like "a fox in a hen house", says Carl Woese, an evolutionary biologist at the University of Illinois in Urbana-Champaign.

Some biologists believe that horizontal gene transfer makes LUCA unknowable⁸.

"Four billion years is plenty of time to scramble the phylogenetic record," says evolutionary biologist Ford Doolittle of Dalhousie University in Halifax, Canada. "Life had achieved its modern cellular status much earlier than anything we can trace back."

Olsen counters that, while trees vary from gene to gene, the average tree that emerges from comparing many genes at once is consistent with the results obtained for ribosomal RNA alone. In other words he believes the phylogenetic record has been eroded, but not erased.

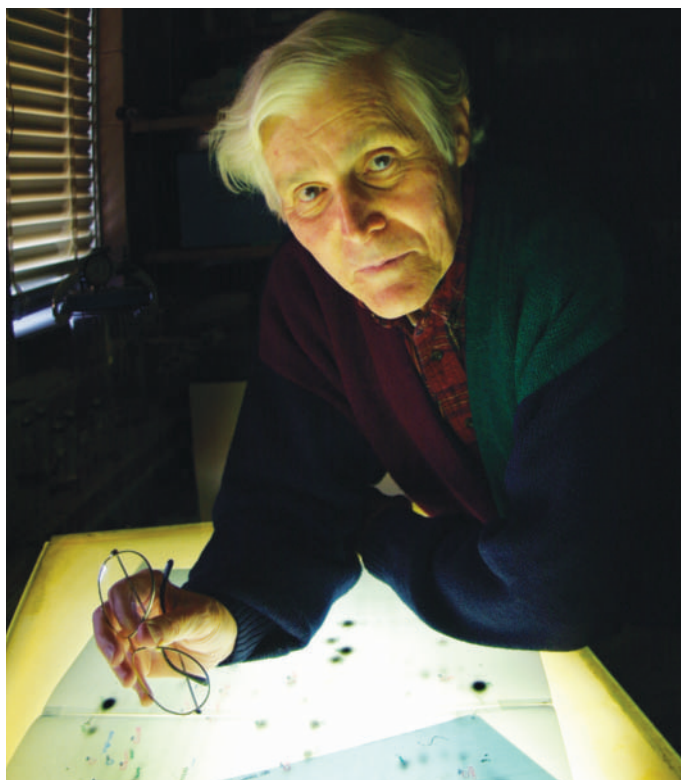
In 1998, the puzzles surrounding the nature of the first life forms led Woese to propose that the universal common ancestor was actually a community of organisms sharing genes⁹. In this communal world, the various primordial organisms had independently come up with solutions to similar problems, such as how to build and expand a membrane, or how to convert organic molecules into useful energy. All of the genes encoding these solutions were available to all the cells in the commune. "I picture genes and their products flowing through a sea of cells," Woese says.

Plug-and-play genes

The genes in this early world would have been 'modular' — their products able to function on their own. Whereas today many cellular functions, such as DNA translation and replication, rely on complex machines encoded in several genes, nearly all the genes in LUCA's community would have been able to function by themselves, like cassettes that can be loaded, removed and replaced. Antibiotic-resistance genes are like that today. A bacterium need only acquire one to fend off an antibiotic.

Ultimately, around 3.5 billion years ago, the modern domains of life would have emerged from the gene-swapping mêlée with many of the genes from the last common community riding on their coat-tails. Inheritance and mutation would then have replaced gene transfer as the most important source of biological novelty as cells became more complex and their functions became less interchangeable. This point, says Woese, was the true origin of species, and so he has christened it the darwinian threshold¹⁰.

Woese's genetic and biochemical commune has become a leading theory of early life, but it has not entirely quashed the concept of LUCA. Koonin, for instance, is still working to provide LUCA with a coherent



Carl Woese: "I picture genes and their products flowing through a sea of cells"

genome by expanding the list of genes it contained beyond the universal 60. He believes modern genes that are almost universal probably also date back to LUCA, and have simply been replaced in a few organisms. So these should be added to the list. It's trickier to tell with less-common genes, and bioinformatic techniques are not yet sophisticated enough to construct a full gene set for LUCA, Koonin says. But it will get easier as sequencing data continue to accumulate.

"The ultimate goal is to arrive at a most defensible reconstruction in terms of a set of genes," he says. This set will number about 600, Koonin estimates, based on what contemporary genomes tell us about the minimum number of genes needed by a self-sufficient organism⁶.

Building a microbe

Once that set of genes is known, it might even be possible to create LUCA in a dish. Building a microbe may seem outlandish, but just such a project is already under way. In 2002, human genome sequencer Craig Venter announced his plan to build an artificial cell with a minimal genome based on modern genes at his Institute for Biological Energy Alternatives in Rockville, Maryland. And last year, a team led by Steven Benner of the University of Florida, Gainesville, used phylogenetic analysis to resurrect a protein from an ancient bacterium that lived around a billion years ago¹¹.

If an artificial LUCA could survive in a laboratory dish, it might deal a blow to the notion of a communal ancestor that depended for

survival on neighbours exchanging genes. That would not surprise Patrick Forterre of the Paris-Sud University in Orsay and the Pasteur Institute in Paris, who says the communal LUCA notion doesn't fit with the way evolution works. "To think of LUCA in terms of a community is to remove the idea of darwinism from early evolution," he says. Although LUCA undoubtedly swapped genes with its neighbours, Forterre argues that it would also have competed with them and ultimately triumphed through some key innovation.

Late last year, Canadian mathematicians provided support for Forterre's case with a model suggesting that a 'commune' of protocells could not persist. Peter Antonelli and Solange Rutz at the University of Alberta in Edmonton wrote equations describing the competition for resources and the sharing of genes and biochemicals that would have gone on in

such a world, and found them to be mathematically unstable. In other words, they say, the commune would have fallen apart¹².

But Woese believes the critique says nothing about his hypothesis. The world of a communal LUCA has so far only been described in qualitative terms, he points out, so there are many mathematical models that could produce an approximate version of it. Also, the modern microbial world is full of bacterial communities that share resources, if not genes.

Of course, finding LUCA would not solve the puzzle of how life began. The idea of a last common community, with a communally sophisticated biochemistry, raises another question: how did all this evolve? This is for someone else to answer, says Woese. "We don't understand how to create novelty from scratch — that's a question for biologists of the future."

John Whitfield is a freelance science writer based in London.

1. Lazcano, A. & Forterre, P. J. *Mol. Evol.* **49**, 411–412 (1999).
2. Nisbet, E. G. & Sleep, N. H. *Nature* **409**, 1083–1091 (2001).
3. Woese, C. R. & Fox, G. E. *Proc. Natl Acad. Sci. USA* **74**, 5088–5090 (1977).
4. Brochier, C. & Philippe, H. *Nature* **417**, 244 (2002).
5. Galtier, N., Tourasse, N. & Gouy, M. *Science* **283**, 220–221 (1999).
6. Koonin, E. V. *Nature Rev. Microbiol.* **1**, 127–136 (2003).
7. Brown, J. R. & Doolittle, W. F. *Microbiol. Mol. Biol. Rev.* **61**, 456–502 (1997).
8. Doolittle, W. F. *Curr. Opin. Struct. Biol.* **10**, 355–358 (2000).
9. Woese, C. *Proc. Natl Acad. Sci. USA* **95**, 6854–6859 (1998).
10. Woese, C. R. *Proc. Natl Acad. Sci. USA* **99**, 8742–8747 (2002).
11. Eric, A. Gaucher, E. A., Thomson, J. M., Burgan, M. F. & Benner, S. A. *Nature* **425**, 285–288 (2003).
12. Antonelli, P. L., Bevilacqua, L. & Rutz, S. F. *Nonlin. Analysis: Real World Appl.* **4**, 743–753 (2003).

Dropping *habilitation* would aid progress in Poland

Standards are high in many disciplines, but changes are still needed in career structure.

Sir—I enjoyed your Editorial “Eastern promise” (*Nature* 426, 369; 2003), but I disagree with Cezary Wójcik’s response, “Eastern Europe; progress stifled by the old guard” (*Nature* 427, 196; 2004). If the system in Poland is—as described by the author—“hierarchical, immobile, hermetic and gerontocratic”, how has it managed to educate postdocs and students who are successful in the best laboratories in the world? The answer is simple: the picture is not as homogeneous as Wójcik suggests. At almost every university or institute of the Polish Academy of Sciences there are scientists who have chosen to remain in Poland, even though they could find positions in the United States or Western Europe.

Disciplines do vary, but in physics, mathematics, biology and chemistry at least, the standards are much higher than Wójcik describes. I am a professor of molecular biology and I probably belong to the “scientific establishment” Wójcik derides. Only three of some 80 publications that I have authored appeared in Polish journals, and I became a professor in Poland after my Polish team published a

paper in *Cell* and several others in *EMBO Journal* and *Proceedings of the National Academy of Sciences*.

I also disagree with the view that “[m]ost research money is distributed by arbitrary administrative decisions, not as peer-reviewed grants”. More than 10 years ago the State Committee for Scientific Research, composed entirely of elected members, introduced a grant system, based on a peer-review procedure that mirrors the US National Science Foundation. All grant applications in the life sciences have to be submitted in English and we ask reviewers from abroad to rank them.

Four years ago we introduced a new system for distributing science funding based only on the quality of the scientific work performed in Polish institutions (see www.kbn.gov.pl). We initiated special programmes on genomics, proteomics, bioinformatics and biodiversity, and we ask scientists from Germany, France, Sweden and the United Kingdom to validate them.

We are collaborating closely with the European Molecular Biology Organization and the Howard Hughes Medical Institute for additional support of young Polish

investigators, who are selected by these distinguished organizations. We are opening new Max Planck–Polish Academy of Sciences laboratories for young scientists working in molecular biology. Similarly, young scientists working at the International Institute of Molecular and Cell Biology in Warsaw are recruited according to international standards (*Nature* 421, 471–472; 2003). Independent positions are awarded on scientific criteria alone, and neither a *habilitation* nor the title of professor is required.

Certainly, change is happening too slowly, especially for young people. Therefore, I am convinced that upon entering the European Union, we have to forget about *habilitations* and introduce international competition for all independent positions, from assistant to full professor.

Maciej Zylicz

International Institute of Molecular and Cell Biology, 4 Ks. Trojdena Street, 02-109 Warsaw, Poland and Chairman of the Biology, Earth Sciences and Environmental Protection Unit, State Committee for Scientific Research, 113 Wspolna Street, 00-529 Warsaw 53, Poland

Polish journals have an international impact

Sir—Cezary Wójcik, in Correspondence (“Eastern Europe; progress stifled the old guard” *Nature* 427, 196; 2004), expresses a view of Polish science that is true only to some extent. As young Polish scientists, working both in Poland and abroad, we wish to offer a counter-balancing viewpoint.

First, we do not deny that in some cases a scientific career may be based on “personal or political connections”. But to suggest this applies to the whole of Polish science, as Wójcik does, is an exaggerated generalization. The *habilitation* is not a mysterious qualification unrelated to any real scientific achievements, but an academic grade having its equivalents in other Western European countries—although in Germany it is no longer required to become a full professor.

Second, the medical sciences require particularly expensive materials and methods, and should not be considered as representative of all scientific disciplines, but rather as an exception.

Similarly, Polish medical journals may not always be internationally recognized, but other disciplines fare better. For example, *Acta Palaeontologica Polonica* has

an impact factor between 0.67 and 1.0 (data for 2001–02), which places it between tenth and fifteenth in the world rankings of the 29 indexed scientific journals in this field, and *Acta Astronomica* (impact factor 3.2 in 2002) ranks ninth in astronomy.

Adam T. Halamski*, Dorota Religa†

**Institute of Paleobiology, Polish Academy of Sciences, Twarda 51/55, 00-818 Warsaw, Poland and Université Claude-Bernard-Lyon 1, UFR des Sciences de la Terre, UMR PEPS, 43 boulevard du 11 novembre, 69 622 Villeurbanne cedex, France*
†Karolinska Institutet, Neurotec, Novum, plan 4, 141 86 Stockholm, Sweden and Medical Research Centre, Polish Academy of Sciences, Pawinskiego 5, 02-106 Warsaw, Poland

Learning humility from a Nobel prizewinner

Sir—The Dutch biologist Niko Tinbergen—whose biography, *Niko's Nature* by Hans Kruuk, is reviewed by John Krebs (*Nature* 427, 293–294; 2004)—may help us to understand one of the problems with German science outlined in the Editorial in the same issue (*Nature* 427, 271; 2004).

Tinbergen had an extraordinary characteristic that accounts, I believe,

for his ability to develop intellectually throughout his life, and stimulate his students and colleagues to do so too. It was his humility. At the weekly seminars that he held in his home for his fellow ethologists and others, I observed that he never hesitated to display his initial incomprehension of a new idea; he was determined to understand. It was an object lesson to his group.

This characteristic was unfortunately not shared by many of his peers. At the European conferences that he organized with his fellow Nobel prizewinners, it was striking how many German professors used their authority to determine what their juniors contributed. This must have been stultifying for both.

Although scientific hierarchies have become less rigid in Germany, Chancellor Gerhard Schröder would do well to encourage more egalitarian intellectual innovation as he develops his plans for several new ‘elite’ German universities.

John Godfrey

41 Lawford Road, London NW5 2LG, UK

correspondence

Contributions to Correspondence may be submitted to corres@nature.com. They should be no longer than 500 words, and ideally shorter. Published contributions are edited.

Invention and innovation

Are novel actions important in the evolution of behaviour?

Animal Innovations

edited by Simon M. Reader & Kevin N. Laland

Oxford University Press: 2003. 344 pp.
£55, \$85 (hbk); £19.99, \$34.50 (pbk)

W. C. McGrew

Innovation is the raw material of evolution, whether it be genetic mutation or the invention of behaviours. No one doubts the former's role in life on Earth through organic evolution, but many doubt that the latter occurs outside human cultural evolution. This book addresses this latter problem, as it is about the "psychological bases, natural ecology, evolution, and adaptive function of behavioral innovation in animals".

The editors — two biologists with research experience ranging from fish to primates — have assembled a volume of 15 chapters based on a symposium held at the 2001 International Ethological Congress in Tübingen, Germany. The authors of the chapters range across animal behaviour, from laboratory-based comparative psychologists such as Hilary Box to fieldworking ethologists (Richard Byrne, for instance). They cover taxa from fish (in the case of Kevin Laland) to great apes (Anne Russon). Not surprisingly, given their big brains and well-studied habits, primates are the focus of seven of the chapters.

Reader and Laland believe that innovation by non-human species is widespread, important and distinctive. In a provocative 35-page introduction, they pose some key questions. Who invents new behavioural patterns? What ecological variables influence innovation? What psychological processes underlie it? And does innovation drive evolution?

Unfortunately, the answers are not straightforward. Novelty is often not clear. Innovation can be idiosyncratic (that is, it never spreads beyond the inventor). Broad definitions equate initiative (for example, the first guppy to learn a new route through a maze) with creativity (such as the composition of new songs by humpback whales). Narrow definitions exclude key elements; for example, if we insist on only totally new patterns, we may exclude interesting modifications of existing behaviour. Innovations can range from mindless serendipity to clever insights. Given these issues, it is not surprising that the editors sometimes contradict themselves and the authors use different definitions for the phenomenon.

One of the chapters, by Hans Kummer and Jane Goodall, is a reprint of an article



A thirst for the new? Blue tits in Britain learnt from one another how to open milk bottles.

published nearly 20 years ago (*Phil. Trans. R. Soc. Lond.* **308**, 203–214; 1985), included presumably because of its landmark status in the field. Its make-up provides a starting point and gives us a benchmark: three pages of pregnant theory from Kummer, followed by nine pages of compelling storytelling about chimpanzees by Goodall. The remainder of the book shows how far the study of innovation has come in the past 20 years.

Pride of place goes to systematic, empirical, comparative analyses that seek to correlate 'innovativeness' with basic key variables. There are reassuring convergences: both Louis Lefebvre and Johan Bolhuis for birds, and Reader and Katherine MacDonald for primates, show a positive correlation between innovation frequency and the size of the neocortex, even when other possibly confounding variables have been controlled for. Byrne shows that much of the deception in non-human primates, especially the novel use of existing behavioural patterns out of context, may be innovative. Russon provides an astonishing array of examples of innovation by free-ranging orang-utans that are being rehabilitated from captivity. Many relate to human activities to which they may have been exposed, but many more cannot be, as humans do not do such things as disassemble palm crowns by hand.

Three chapters are given over to birds, which almost always provide bigger and better data sets than do mammals. Russell Greenberg contends that a preference for the familiar over the novel provides the key distinction that makes sense in terms of ecological analyses. Peter Slater and Robert Lachlan focus on bird song and distinguish

among four alternative sources of novelty: immigration (new-comers and diffusion), innovation (modification of existing elements), invention (really new and underived) and improvisation (extemporaneous and transient). The best experimental work continues to be done on birds, a point that blinkered primatologists sometimes fail to acknowledge.

What to make of the claims of innovation by reptiles and fish? Laland and his students have produced a series of elegant experiments with guppies which may well exceed in quality of design any comparable studies of primates, for obvious logistical reasons. However, to be frank, it has yet to be shown that guppies do anything interesting, and if solving a simple maze task counts as innovation, it is hard to know how to compare it meaningfully with, say, elementary technology in apes or the riff-like singing of song birds, to quote Marc Hauser.

The editors have succeeded in putting innovation well and truly on the map as a phenomenon to be reckoned with. They provide a series of outstanding questions that demand answers, and the book will stimulate further research efforts. In that sense, Reader and Laland's volume can be compared to Benjamin Beck's *Animal Tool Behaviour* (Garland, 1980) or *Machiavellian Deception* by Richard Byrne and Andrew Whiten (Oxford University Press, 1988), two seminal tracts that set high standards.

W. C. McGrew is professor of anthropology and zoology at Miami University, Oxford, Ohio 45056, USA. His latest book, *The Cultured Chimpanzee: Reflections on Cultural Primatology* (Cambridge University Press), will be published later this year.

Synthetic thought

Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence

by Pamela McCorduck

A. K. Peters: 2004. 576 pp. \$19.95

John L. Casti

For centuries humans have speculated on how the technology of the time could be used to carry out mental tasks that mimic, if not surpass, those done by the human mind. The current incarnation of this form of hubris landed on the intellectual landscape in a legendary 1950 paper in which British computer pioneer Alan Turing laid out the agenda for the creation of a "machine that thinks" (*Mind* 59, 433–460). Just six years later, a summer conference at Dartmouth College in New Hampshire brought together an eclectic group of maverick mathematicians, engineers, psychologists, computer scientists (in today's terminology), neurobiologists and other assorted denizens of the academic world. They established a research programme in what one attendee, John McCarthy, dubbed "artificial intelligence" (AI), an evocative (and provocative) appellation by which the field has been known ever since.

Like all zealots setting up a new religion, the Dartmouth group asserted some pretty amazing claims. Two of the most interesting to catch the public eye ended up as touchstone problems by which progress in AI could be measured. The first was to develop within ten years a computer program that could beat the world chess champion. The second was to develop in about the same length of time a computer program that could translate from one human language to another at a level indistinguishable from that of a professional translator. The rationale for these benchmark problems was that, in both cases, creating a program to carry out the task would teach us many things about the mysterious ways of the human mind.

Amusingly, the first goal has now been achieved by the program Deep Blue II — but it taught us nothing about human thought processes, other than that world-class chess-playing can be done in ways completely alien to the way in which human grandmasters do it. The second goal, however, is about as far from being achieved as the number of atoms in the Universe is from infinity. But, strangely perhaps, work on machine language translation has taught us a lot about how human language processing takes place.

About 20 years after the Dartmouth meeting, writer Pamela McCorduck, wife of pioneering computer scientist Joseph Traub, was having lunch at Stanford with two AI veterans, McCarthy and Edward Feigenbaum. She suggested getting the thoughts



Mind games: world chess champion Garry Kasparov met his match in the computer Deep Blue.

and impressions of the workers in this field down on paper as a kind of historical and sociological account of the emergence and evolution of an entirely new field of intellectual endeavour.

In 1979, after several years of interviews and lunchtime conversations with AI researchers, including Marvin Minsky, Lotfi Zadeh, Herbert Simon and Alan Newell, McCorduck published her book, *Machines Who Think*. This work did not pretend to be an exhaustive account of the entire field, but rather was a kind of eclectic sampling of various schools of thought in AI and how much progress, or lack thereof, had been made towards the ultimate goal of a thinking machine.

Being an informed outsider to the field, as well as a gifted writer and storyteller, was a decided advantage to McCorduck. She had no particular axe to grind with regard to championing one approach over another, and was able to tell the human story of AI in terms that made the book fascinating reading — even if you were not especially interested in whether a machine could one day replace your brain.

In retrospect, 1979 was a turning point for AI. The technological advances of the preceding two decades, coupled with the appearance that year of Douglas Hofstadter's

Pulitzer-prize-winning book *Goedel, Escher, Bach*, catalysed a resurgence of 'bottom-up' AI based on neural networks, or what came to be termed 'connectionism' — an approach that had long been displaced by 'top-down' designed programs. Moreover, in the decades since, much progress has been made in robotics, distributed intelligence, cooperative intelligence and numerous other areas, to go along with an exponential improvement in hardware, which allowed these approaches to be implemented.

McCorduck has now reissued her original 1979 book, this time with an afterword of more than 100 pages that tells the story of the past 25 years. If you are interested in how the pioneers of AI approached the problem of getting a machine to think like a human — a story told here with verve, wit, intelligence and perception — there is no better place to go than this book. It is a worthy continuation of the story begun in McCorduck's 1979 account. One can only hope that she will have the same stamina and enthusiasm to produce an expanded, expanded edition for our entertainment and enlightenment in another decade or two. No one could do it better.

John L. Casti is at Complexica, Santa Fe, New Mexico 87505, USA, and the Institute for Monetary Economics, Vienna, Austria.



Art

Go with the floe

During his three-month sojourn with the British Antarctic Survey, artist John Kelly experienced the polar region as sublime but often darkly ominous. This painted-over photomontage — part of a six-part series called *Silent Sea* —

portrays the Antarctic's threatening beauty. It is part of an exhibition of his work, "Due South: Art and the Antarctic", which runs from 24 February to 1 August at the Natural History Museum in London.

Alison Abbott

A recipe for the mind

The Birth of the Mind: How a Tiny Number of Genes Creates the Complexities of Human Thought

by Gary Marcus

Basic Books: 2004. 240 pp. \$26, £19.99

Anthony P. Monaco

If the mind can be explained from the workings of the brain, and the brain develops by direction from our genes, then presumably the mind can be explained from our genetic make-up. But how can only 30,000 genes make a brain with billions of neurons and encode the particular aspects of cognition that make us human?

The Birth of the Mind tries to unravel this complex problem by first explaining what we know about each component of the argument: the mind, the brain, our genes and the environment. The breadth of examples used to achieve this is impressive, encompassing 40 different organisms (from bacteria to chimpanzees), 30 different genes and 20 different brain regions.

The author, Gary Marcus, spends much of his efforts building up the reader's knowledge base. It is difficult to make an argument that involves such diverse disciplines as evolution, genetics, gene expression, cell biology, neurobiology and psychology without teaching the reader the bare essentials. Marcus does particularly well to make the relevant issues in these areas understandable

to the lay reader, and does an even better job of dispelling the myths that impede the way we think about genes and their role in making brains, and hence minds.

Marcus is a cognitive psychologist who understands genes. He has researched his topic well and describes the complex world of genes in an entertaining and gripping way. He dispels the myth that there are too few genes by explaining that single genes can encode several proteins with different functions, and more importantly that genes can be turned on and off in groups in multiple combinations to perform highly orchestrated and complex functions. He discards the analogy of genes as blueprints for building a brain (or any other organ in the body), and prefers to think of genes as the 'recipe' required for the correct development of the brain. He explains heritability and the difference between single-gene effects — and their resulting monogenic and relatively rare diseases — and complex genetic interactions, which cause more common diseases.

He also unravels the paradox of flexibility. How does the brain of a newborn, with its complex structures and connections, have the plasticity to enable it to respond to environmental influences as it develops further? Marcus disentangles the nature-versus-nurture argument using many examples from neuroscience research that show that the brain is built by genes in a self-organized way before being reorganized and shaped by experience and the environment. It is not a battle where one side wins, but a vital interaction.

Having clarified these two paradoxes using our current knowledge of genetics and

neuroscience, can we explain how genes make minds? The story is only beginning. This book shows that genes build brains and that brains are designed to be flexible and to learn, but the jump from genes to the mind is an indirect one. The question cannot yet be answered, and it is not entirely clear where the answer will come from.

Will geneticists pinpoint genes and gene pathways that will inform us about human cognition? This is already happening but will provide only part of the answer and confirm that genes are directly influencing mental abilities. Will cognitive scientists detail the relationship between neural structures and mental structures? This area of research is expanding, thanks to modern brain-imaging techniques that enable researchers to 'see' the brain in action. The path ahead to integrate these disciplines to gain a fuller understanding is optimistically vague, and anyone interested in the topic would be encouraged to read this book.

Lay readers may be daunted by the sheer complexity of the science — even with the best intentions of explanations in lay terms, a glossary and an appendix to explain the genes. But the more scientifically minded, especially those with a background in either genetics or neuroscience, would gain much from the book. It would at least dispel some myths and paradoxes, leaving the possibilities open for an eventual understanding of how 30,000 genes can provide the recipe for the mind.

Anthony P. Monaco is director of the Wellcome Trust Centre for Human Genetics, University of Oxford, Headington, Oxford OX3 7BN, UK.

One more thing...

How an extra control experiment led to a change of research field.

Nancy Rothwell

The excitement of an important scientific result and the satisfaction of preparing the data for publication is often tempered by critical analysis of what is missing in the story. There always seems to be that one extra control experiment that you know has to be done, not least because the referees will ask for it after you submit the paper. It's the one that will be uninteresting and time consuming, and is often left until last. It was one such control experiment that completely changed the direction of my research.

I trained as a physiologist, and my early research career investigated the mechanisms of energy balance regulation in mammals, the role of thermogenesis in energy homeostasis and the disruptions that lead to obesity. Moving north to the University of Manchester in the late 1980s meant establishing an independent research group for the first time. I had already conducted some studies on cachexia (the wasting condition that accompanies severe disease and injury) and on the role of cytokines such as tumour necrosis factor- α and interleukin-1 (IL-1) — although immunology wasn't really my cup of tea, I found these molecules fascinating.

The University of Manchester had a strong reputation for research on experimental and clinical aspects of injury, so I set up my modest lab and embarked on establishing collaborations with relevant scientists and clinicians, with the goal of discovering the causes of cachexia — and perhaps even identifying new treatments (ambitious in hindsight, but not at the time).

One of the first of these collaborations was with a recently appointed lecturer, Celestine O'Shaughnessy, and began — like so many fruitful collaborations — because we had adjacent offices and got on well together. Celestine's background was in neuroscience, which was then a mystery to me. She was working on experimental cerebral ischaemia (stroke) in rats, an area very different to my own. But it was well known from clinical studies that brain damage causes a marked hypermetabolic state and often leads to cachexia. So we got together — a good collaboration of complementary expertise — to ask if ischaemia does induce hypermetabolism, and how.



Controls and career tracks: where will that extra experiment lead you?

The work went well and we published several papers demonstrating a significant increase in metabolic rate and accompanying hyperthermia in response to cerebral ischaemia, which was dependent on activation of the sympathetic nervous system. Celestine took on a new PhD student, Jane Relton, and we planned further studies to specifically investigate the role of IL-1 in responses to cerebral ischaemia, and its actions in the brain. But events took a different turn when Celestine was seduced by an attractive position in the pharmaceutical industry. I was left without my neuroscience expert, but with a newly acquired PhD student.

Fortunately, Jane was experienced in working on the brain, and completed many experiments, showing that blocking of endogenous IL-1 by administering a naturally occurring IL-1 receptor antagonist (IL-1ra) to rats prevented the increases in metabolic rate induced by a stroke. Others in my research group showed similar effects of cytokine modification in animals exposed to peripheral infection or injury, and we identified several downstream mediators of IL-1 action. All was going well.

As we were writing up Jane's experiments on IL-1, we knew that there was an extra control experiment that needed to be done — to check that IL-1ra was specifically influencing hypermetabolism, and not directly altering

the damage caused by ischaemia. This was a time-consuming study, and of course we knew the outcome — IL-1 could not contribute to injury within the brain, which was believed to be barely influenced by events or mediators of the immune system (the brain was considered to be an 'immune-privileged organ'). But we were surprised by the results. IL-1ra reduced the damage by well over 50%, and did so consistently. Still barely believing the results, we completed further experiments and published the work, but it raised little attention at the time — probably no one else believed it either.

This presented me with a dilemma. I believed (even if not many others did) that this was an important observation, which challenged the prevailing dogma of the brain as an immune-privileged organ. The problem was how could I, as a newly independent scientist with little more than undergraduate knowledge of neuroscience, take forward a project on cerebral ischaemia?

I was still young enough to believe that everything is possible and that I should follow what my instinct told me was an important finding. Jane continued with her experiments, and others in the lab diverted to IL-1 and stroke, while I avidly read the neuroscience literature and went to every meeting on stroke-related research that I could. I was reassured by colleagues in industry who said that you can change fields and become re-established in quite a short time. But this was a difficult decision. I was established in the field of metabolism, had made some inroads with immunology, but had little experience in the field of neuroscience.

Enthusiasm prevailed over logic and we launched into research on stroke and, just over a decade later, we have completed the first small clinical trial of IL-1ra in stroke and are beginning to uncover the mechanisms of IL-1 induction and action in the brain. But I have not abandoned my interest in metabolism entirely, and I no longer know what to call myself. Am I a physiologist, neuroscientist, immunologist or endocrinologist?

This experience taught me two things: do the important control experiment early, and follow your intuition and enthusiasm — even if it's wrong it will be more fun. ■

Nancy Rothwell is at the School of Biological Sciences, University of Manchester, Manchester M13 9PT, UK.

Immigration denied

Pasko Rakic

The adult human brain cannot replace lost neurons. This might be because it is reluctant to accept newcomers into an already established neural network, rather than because potential progenitors are absent.

Neural stem cells are a focus of strong interest because of the possibility that they could be used to replace neurons that have been damaged or lost — perhaps as a result of injury such as trauma or stroke, or through neurodegenerative disorders such as Parkinson's disease. These stem cells can give rise to neurons and their supporting cells, glia, and it is hoped that something akin to neural stem cells in the adult human brain could be stimulated to generate replacement neurons.

Non-mammalian vertebrates, such as the salamander, can regenerate large portions of their brain and spinal cord, but humans have evidently lost this capacity during evolution. Therefore, most research on neural stem cells is carried out on mammals such as rodents, which are genetically closer to humans. However, although mammalian genomes may be similar, this similarity masks vast species differences in the way the brain is organized and in its capacity for regeneration and susceptibility to environmental insults. The failure of brain repair in clinical trials based on the promising results seen after the use of similar procedures in rodents is sobering testimony to the importance of such species-specific distinctions.

Human neural stem cells behave differently from their rodent equivalents in culture¹, but the direct study of human brain tissue by Sanai *et al.*², described on page 740 of this issue, shows additional significant and clinically relevant species-specific differences. The authors' investigations on a large number of postmortem and biopsy samples reveal two basic findings. First, neural stem cells that can potentially give rise to neurons, as well as to two types of glial cell (astrocytes and oligodendrocytes), are situated in a region of the forebrain known as the subventricular zone. Second, a pathway known as the rostral migratory stream — which in adult rodents contains neurons that migrate from the subventricular zone to the brain region concerned with sensing smell — is absent in humans.

In adult mammals, including humans, the subventricular zone (more commonly known as the subependymal zone^{3–5}) contains cells that have the characteristics of glial cells and that can generate neuronal cells in culture^{6–8}. Sanai *et al.*² show that in adult humans these 'glial progenitor cells' form a prominent layer, or ribbon, that is restricted

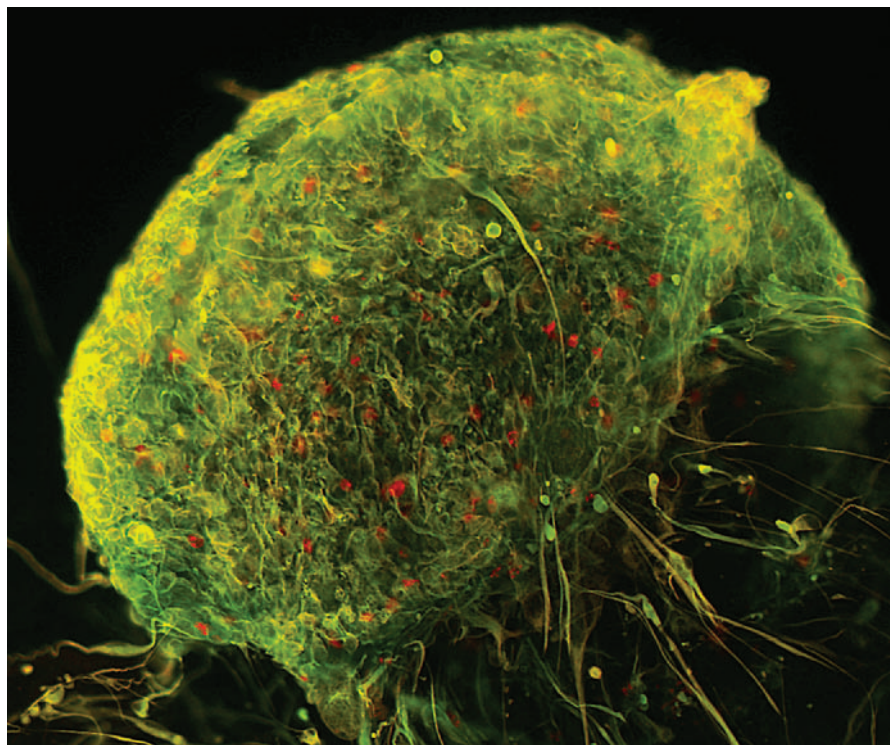


Figure 1 Proliferative potential. Astrocytes that can behave as neural stem cells are present in the walls of the lateral ventricle in the adult human brain. This culture containing proliferating astrocytes labelled with glia fibrillary acidic protein (GFAP; green) and the cell division marker Ki-67 (red) was derived from cells taken from a patient during the removal of part of the walls of the lateral ventricle².

to a specific region in the brain that lines the lateral cerebral ventricle (Fig. 1). This region is also present in non-human primates, but it is thinner and less well delineated than in humans^{4,9–11}. Based on the knowledge that about 0.7% of cells in the human subventricular zone contain a specific nuclear protein, Ki-67, that is associated with DNA synthesis and hence cell division, Sanai and colleagues² infer that these cells can proliferate. However, the number of cells that are actually dividing is probably smaller than this, as Ki-67 can also be present in dying cells^{12,13}. Importantly, although the cells that were positive for Ki-67 also contained marker molecules for astrocytes, they did not contain markers for immature neurons. This observation is similar to that seen with subependymal cells in other parts of the adult mammalian brain, such as the spinal cord or brain stem, which also do not generate neurons *in vivo*. However, if these cells are cultured or are transplanted to acceptable cellular environments such

as the hippocampus, they can show some neuronal features¹⁴.

Sanai and colleagues' second key finding² is the absence in humans of the rostral migratory stream, which in mammals with a highly developed sense of smell supplies neurons to the olfactory bulb. It seems that these migratory neuronal precursors are present in humans during infancy but disappear during childhood⁵. This is not unexpected, considering the dramatic loss of olfactory receptor genes in the relatively small human olfactory bulb¹⁵ and the substantial reduction in the size of the rostral migratory stream in non-human primates¹⁶. However, it highlights the danger of extrapolating results from rodents to humans. Sanai and colleagues' findings therefore add to the evidence that, with the exception of a specific cell type in the hippocampus¹⁷, the human brain shows no sign of spontaneous neuronal turnover^{11,18}.

Back to the potential glial progenitors, then — what might be their function, given that they do not give rise to neurons? Apart

N. SANAI *ET AL.*

from serving as a major source for the turnover of astrocytes and oligodendrocytes, which occurs normally in the adult primate brain, they also provide a reservoir for a certain type of astrocyte that migrates to sites of traumatic, infectious or degenerative brain damage. In addition, their uncontrolled proliferation is considered to be a main cause of certain types of invasive glial brain tumour. However, it is instructive that these dormant glial progenitors never generate malignant tumours of any neuronal cell type. This indicates that there are probably formidable suppressors that block the production of neuronal cells in the adult human brain¹¹.

The main lesson from this comprehensive study is that the lack of neuronal turnover and/or replacement of injured neurons in the adult human brain is not due to the absence of potential neural stem cells. Rather, it is more likely to be due to a remarkable resistance to accept such cells into a mature neuronal network. The systematic decrease in the extent of adult neurogenesis during vertebrate evolution, culminating in primates, may be the result of an adaptation to keep neuronal populations with their accumulated experience for an entire lifespan¹¹. Although the therapeutic transplantation of new neurons to regions of the human brain that are responsible for more advanced brain functions may be counterproductive, their transplantation to other regions, such as the sensory or motor systems of the brain, could have enormous clinical significance. So the identification of the cellular and molecular mechanisms that prevent adult neural stem cells from becoming integrated into functional neuronal networks would be a major achievement. Without this, the stimulation of dormant neural stem cells, and even their survival, may not be sufficient to accomplish brain repair. ■

Pasko Rakic is in the Department of Neurobiology, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06511, USA.
e-mail: pasko.rakic@yale.edu

- Ginis, I. & Rao, M. S. *Exp. Neurol.* **184**, 61–77 (2003).
- Sana, N. et al. *Nature* **427**, 740–744 (2004).
- Lewis, P. D. *Nature* **217**, 974–975 (1968).
- McDermott, K. W. & Lantos, P. L. *Brain Res. Dev. Brain Res.* **57**, 269–277 (1990).
- Weickert, C. S. et al. *J. Comp. Neurol.* **423**, 359–372 (2000).
- Kirschenbaum, B. et al. *Cereb. Cortex* **4**, 576–589 (1994).
- Laywell, E. D., Rakic, P., Kukekov, V. G., Holland, E. C. & Steindler, D. A. *Proc. Natl Acad. Sci. USA* **97**, 13883–13888 (2000).
- Palmer, T. D. et al. *Nature* **411**, 42–43 (2001).
- Kornack, D. R. & Rakic, P. *Science* **294**, 2127–2130 (2001).
- Koketsu, D., Mikami, A., Miyamoto, Y. & Hisatsune, T. *J. Neurosci.* **23**, 937–942 (2003).
- Rakic, P. *Nature Rev. Neurosci.* **3**, 65–71 (2002).
- Yang, Y., Geldmacher, D. S. & Herrup, K. J. *Neurosci.* **21**, 2661–2668 (2001).
- Klein, J. A. et al. *Nature* **419**, 367–374 (2002).
- Lie, D. C. et al. *J. Neurosci.* **22**, 6639–6649 (2002).
- Gilad, Y., Man, O., Pääbo, S. & Lancet, D. *Proc. Natl Acad. Sci. USA* **100**, 3324–3327 (2003).
- Kornack, D. R. & Rakic, P. *Proc. Natl Acad. Sci. USA* **98**, 4752–4757 (2001).
- Eriksson, P. S. et al. *Nature Med.* **11**, 1313–1317 (1998).
- Seress, L., Abraham, H., Tornoczky, T. & Kosztolanyi, G. *Neuroscience* **105**, 831–843 (2001).

Palaeoclimate

Low-down on a rhythmic high

Katharina Billups

Changes in the amount of solar energy reaching Earth account for certain climate cycles at high and low latitudes. Surprisingly, the effects of a high-latitude cycle evidently reached into the tropics.

During the 1920s, Milutin Milankovitch, a Serbian mathematician, calculated the effects of alterations in Earth's motion around the Sun on the amount of solar energy reaching different latitudes¹. Since then, some of the long-period, cyclic changes seen in archives of past environmental conditions on Earth have been explained by these changes in insolation. For example, a 41,000-year cycle in a climate record is commonly ascribed to changes in the planet's tilt, which affects insolation at high latitudes, and a 19,000–23,000-year cycle is ascribed to changes in the planet's wobble, which dominates insolation at low and middle latitudes^{2,3}.

Eighty years after Milankovitch's innovative work, Liu and Herbert (page 720 of this issue⁴) have unearthed surprising evidence that the views based on his conclusions are incomplete. The authors provide a record of low-latitude sea surface temperature (SST) that is in phase with the 41,000-year rhythm reminiscent of high-latitude insolation. Although their record represents only the past 1.8 million years, a small portion of Earth's history, their findings force us to think further about a climate system that we already knew to be complex.

How does the geological record document past climate change? Perhaps the most

ubiquitous recorders of past climate change are the fossil shells of foraminifera, calcareous marine organisms. To a large degree, the oxygen-isotope ratio of the calcium carbonate shell of foraminifera reflects the oxygen-isotope ratio in sea water, which in turn is a function of the hydrologic cycle and thus the amount of fresh water stored as ice at the poles. Because ice-sheet growth and decay are affected by cycles in insolation, the $H_2^{18}O/H_2^{16}O$ ratio of sea water varies at the same frequency as does the $^{18}O/^{16}O$ ratio of the carbonate shells of the foraminifera in the world ocean. Over time, foraminifera accumulating on the bottom of the ocean build an archive of information about glacial to interglacial climate change.

These types of record provide the backbone of palaeoclimate studies. For example, over the past 5 million years perhaps the most profound change in Earth's climate history was the onset of glaciation in the Northern Hemisphere, recorded by the increase in foraminiferal $^{18}O/^{16}O$ ratios between 3 million and 2 million years ago (Fig. 1). However, the isotopic composition of sea water, and hence global ice extent, is not the only variable determining foraminiferal $^{18}O/^{16}O$ ratios; water temperature affects the preference of one isotope over the other incorporated in calcium carbonate during

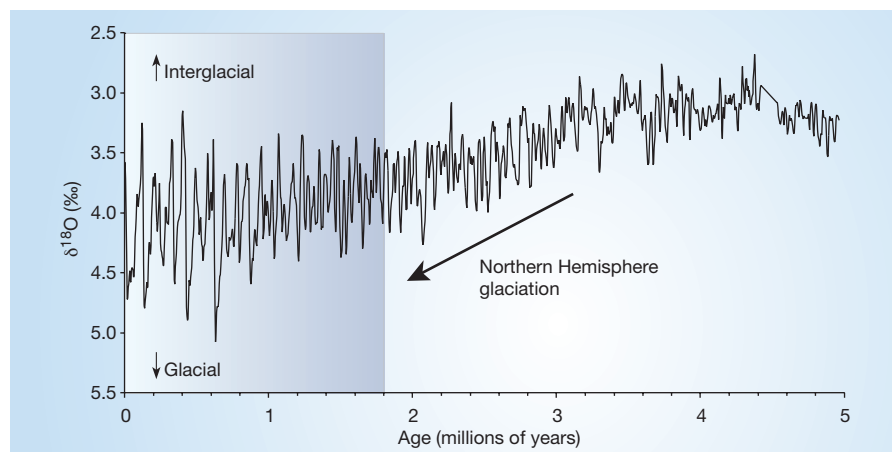


Figure 1 Global climate change from 5 million years ago to the present. This record is inferred from the foraminiferal oxygen-isotope archive from the eastern equatorial Pacific¹², with $^{18}O/^{16}O$ ratios plotted as a per mil deviation from a standard — $\delta^{18}O$ (‰). Over this time there has been a general trend towards cooler climate; at the moment, we are in a peak warm interval within the overall cooling trend. The increase in foraminiferal $\delta^{18}O$ values between 3 million and 2 million years ago signifies the (poorly understood) growth of ice sheets in the Northern Hemisphere. Liu and Herbert⁴ have constructed a separate record for the eastern equatorial Pacific from 1.8 million years ago (shaded), and provide evidence that cyclic processes operating at high latitudes produce a surface-ocean response at low latitudes.

calcification. So although foraminiferal $^{18}\text{O}/^{16}\text{O}$ ratios by themselves do provide information about large-scale climate change as described above, they cannot contribute unequivocally to our understanding of the relationship between individual climate variables such as ice volume and water temperature.

Against the backdrop of global climate change, Liu and Herbert⁴ introduce an SST record from the eastern equatorial Pacific Ocean that spans the past 1.8 million years. The data are derived from measurements of the biochemistry of certain species of marine phytoplankton that are preserved in deep-sea sediments. At the molecular level, the structure of these organisms is a function of temperature⁵. The record generated by Liu and Herbert thus provides a unique opportunity to set SST variations in the eastern equatorial Pacific in the context of global climate change — as, for instance, assessed by the oxygen-isotope composition of foraminifera shown in Fig. 1.

Liu and Herbert are not the first to apply this 'palaeothermometer' to the geological record^{6–8}, but they provide the first tropical SST record to go as far back as 1.8 million years. This record shows strong 41,000-year fluctuations in its early part (1.8–1.2 million years ago). Liu and Herbert observe that the fluctuations correlate closely with variations in foraminiferal $^{18}\text{O}/^{16}\text{O}$ ratios and, hence, global ice volume. This is perhaps a result of feedbacks between ice extent and the amount of insolation being reflected from Earth's surface. But as the authors point out, SST minima just precede ice-volume maxima (and vice versa), which means that the ice sheet cannot be the initial driving mechanism. Liu and Herbert allow that it is surprising that this connection is strongest between 1.8 million and 1.2 million years ago, when Northern Hemisphere ice sheets were still relatively small. Importantly, however, SST minima and maxima in the eastern equatorial Pacific are coincident with minima and maxima in insolation at high latitudes, and thus more suggestive of a coupling between the low and high latitudes via circulation in the atmosphere or upper ocean, or both. Regardless of the mechanism, Liu and Herbert present new evidence for a high-latitude signature in a low-latitude climate record.

These results are particularly intriguing in light of the search for the causes of glaciation in the Northern Hemisphere between 3 million and 2 million years ago (Fig. 1). Current explanations focus on tectonic events that had a large effect on patterns of ocean circulation, such as the closure of the Central American⁹ and Indonesian¹⁰ seaways. It has also been proposed that the onset of glacial–interglacial climate cycles around 3 million years ago should have been accompanied by an increased response of the tropical oceans to changes in insolation caused by

the planet's tilt; this should be recorded as 41,000-year cycles in low-latitude climate records¹¹. Liu and Herbert's SST record from the eastern equatorial Pacific supports this idea. Ultimately, the hypothesis needs to be tested by constructing similar records that include this entire climate transition. But for the moment, Liu and Herbert have provided an enticing example of the strength of the connection between high and low latitudes. ■

Katharina Billups is at the College of Marine Studies, University of Delaware, 700 Pilottown Road, Lewes, Delaware 19958, USA.

e-mail: kbillups@udel.edu

1. Milankovitch, M. *Serb. Akad. Beogr. Spec. Publ.* 132 (1941).
2. Hays, J. D. *et al. Science* **194**, 1121–1132 (1976).
3. Berger, A. L. *J. Atmos. Sci.* **35**, 2362–2367 (1978).
4. Liu, Z. & Herbert, T. D. *Nature* **427**, 720–723 (2004).
5. Wefer, G., Berger, W. H., Bijma, J. & Fischer, G. in *Use of Proxies in Paleoclimatology* (eds Fischer, G. & Wefer, G.) 1–68 (Springer, Berlin, 1999).
6. Emeis, K. C., Dooze, H., Mix, A. & Schulz-Bull, D. *Proc. ODP Sci. Results* **138**, 605–614 (1995).
7. Sikes, E. & Keigwin, L. D. *Paleoceanography* **9**, 31–45 (1994).
8. Sachs, J. P. & Lehman, S. J. *Science* **286**, 756–759 (1999).
9. Driscoll, N. W. & Haug, G. H. *Science* **282**, 436–438 (1998).
10. Cane, M. A. & Molnar, P. *Nature* **411**, 157–162 (2001).
11. Philander, S. G. H. & Fedorov, A. V. *Paleoceanography* **18**, 10.1029/2002PA000837 (2003).
12. Mix, A. *et al. Proc. ODP Sci. Results* **138**, 371–412 (1995).

Transcription

Origins of licensing control

Xue Li and Michael G. Rosenfeld

Organ development requires precise regulation of both the total number and the different types of cells. Much is known about how each process is controlled, but new light has been shed on how the two are linked.

From fertilization to maturity, a single set of genetic information must be accurately passed on to each daughter cell during the cell cycle; the mechanism involved is known as 'replication licensing'. At the same time, specialized cells and organs use unique combinations of genes, which are switched on or off by factors that control gene transcription, and this process — cell differentiation — determines the types of cell that will arise. These two processes are expected to be

tightly integrated during development in order to generate organs of the correct size. On pages 745 and 749 of this issue, Wittbrodt *et al.*¹ and Kessel *et al.*² provide further insights into the molecular interactions between regulation of the cell cycle and regulation of cellular differentiation.

Replication licensing involves the sequential assembly of components of a replication complex onto the regions from which DNA replication begins³. At the heart

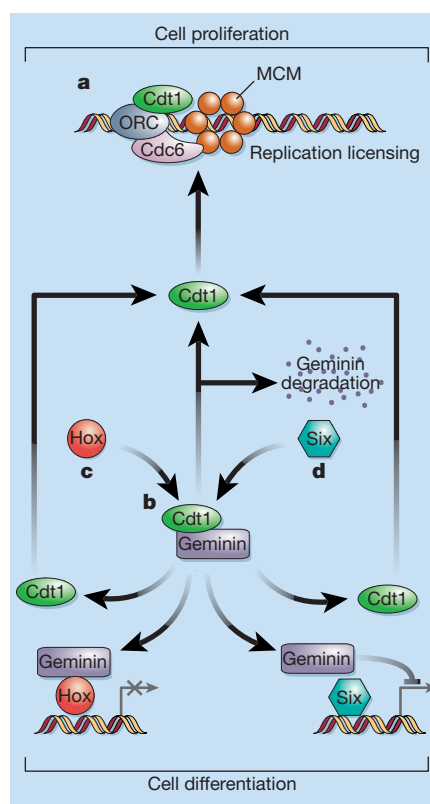


Figure 1 Regulation of cell proliferation and differentiation by Six and Hox proteins.

a, Cdt1 is an essential component of the replication complex, which mediates proliferation. b, The interaction of geminin with Cdt1 inhibits assembly of the minichromosome maintenance complex (MCM), so proliferation cannot occur unless geminin is degraded. c, Kessel and colleagues² have found that Hox can displace geminin from Cdt1, thereby allowing proliferation to proceed. Geminin displaces Hox from its target genes and/or can interact with polycomb proteins to influence Hox activity. d, Wittbrodt and colleagues¹ show that an interaction between Six and geminin can displace geminin from Cdt1 and so induce proliferation. Geminin antagonizes Six3 transcriptional activity but does not interfere with its ability to bind DNA. Geminin regulates cell differentiation by directly interacting with Six and Hox.

of this regulatory machinery are families of proteins that include the 'origin recognition complex' and the 'minichromosome maintenance complex'. But before the latter complex can bind, two other proteins — Cdc6 and Cdt1 — must be in place. And the formation of the replication complex itself is regulated both positively and negatively during the cell cycle so that replication occurs only once. A protein known as geminin is involved in inhibiting replication licensing^{4,5} — it interacts with Cdt1 and so prevents the assembly of the minichromosome maintenance complex (Fig. 1).

Kessel and colleagues² have identified a strong interaction between geminin and transcription factors of the Hox and polycomb protein families, which are essential for embryonic patterning and the specification of tissues and organs^{6,7}. The interaction between Hox and geminin can displace geminin from Cdt1, and so promotes DNA replication and cell proliferation (Fig. 1). Using a series of approaches in chick embryos, Kessel and colleagues found that geminin interacts with polycomb to modify the region in which the *Hoxb9* gene is expressed. The authors also suggest that geminin behaves *in vivo* like a polycomb protein — it inhibits *Hox* gene activation. Because polycomb factors are likely to be recruited in response to specific epigenetic markers that influence gene expression, there may be other components involved in this inhibitory mechanism⁷.

Meanwhile, Wittbrodt and colleagues¹ investigated the functions of the Six family of transcription factors, which are linked to development and proliferation in certain tissues. Six3 and Six6, the most closely related members of the Six gene family in vertebrates, are required for the development of the eye^{8,9}. When Wittbrodt and colleagues searched for factors that interact physically with Six3, they found geminin — so just as geminin can bind Hox, it can also interact with Six3. Six3 seems to promote cell proliferation by displacing geminin from Cdt1 (Fig. 1). Using the retina as a model system, Wittbrodt and colleagues show that inhibiting the function of geminin (loss-of-function), similar to increasing the function of Six3 (gain-of-function), promotes cell proliferation and so increases retina size. Geminin gain-of-function, like Six3 loss-of-function, can be restored by overexpressing Six3. Six3 and Six6 might also control cell proliferation by directly regulating specific genes that control growth^{8,9}. This study, therefore, also provides the first evidence that members of the Six family can function in both a transcription-dependent and -independent manner by regulating replication-licensing events through geminin and Cdt1.

So a role for geminin in controlling cell proliferation is emerging. But how does it control cell differentiation¹⁰? The present

studies suggest that it might do so by interacting directly with genes such as *Six* and *Hox*. In this regard, the new results^{1,2} support different notions. Wittbrodt and colleagues¹ show that geminin antagonizes the function of Six3 in gene transcription without interfering with its DNA-binding activity. On the other hand, Kessel and colleagues² suggest that geminin can antagonize Hox function by displacing Hox from the genes that it targets and/or by regulating *Hox* gene expression by interacting with specific polycomb gene pathways. Perhaps the discrepancy reflects the differences between transcription factors?

So the results of both groups provide an intriguing additional molecular link between the control of cell differentiation and the control of proliferation — they support a general model in which genes that are involved in regulating development also control cell proliferation by directly inhibiting the interaction of geminin with Cdt1. Geminin is present only when the nucleus divides during the cell cycle¹¹, so it will be of particular interest to find out whether this

cell-cycle-specific expression of geminin is involved in determining the cell type. Also, to what extent might this principle apply to other proteins that are similar to Hox and Six, and to other classes of transcription factor? ■

Xue Li and Michael G. Rosenfeld, a Howard Hughes Medical Institute investigator, are in the School and Department of Medicine, University of California, San Diego, 9500 Gilman Drive, CMM-W345, La Jolla, California 92093-0648, USA.

e-mails: seli@ucsd.edu

mrosenfeld@ucsd.edu

1. Del Bene, F., Tessmar-Raible, K. & Wittbrodt, J. *Nature* **427**, 745–749 (2004).
2. Luo, L., Yang, X., Takihara, Y., Knoetgen, H. & Kessel, M. *Nature* **427**, 749–753 (2004).
3. Bell, S. P. & Dutta, A. *Annu. Rev. Biochem.* **71**, 333–374 (2002).
4. Tada, S., Li, A., Majorano, D., Mechali, M. & Blow, J. J. *Nature Cell Biol.* **3**, 107–113 (2001).
5. Wohlschlegel, J. A. et al. *Science* **290**, 2309–2312 (2000).
6. Krumlauf, R. *Cell* **78**, 191–200 (1994).
7. Orlando, V. *Cell* **112**, 599–606 (2003).
8. Li, X., Perissi, V., Liu, F., Rose, D. W. & Rosenfeld, M. G. *Science* **297**, 1180–1183 (2002).
9. Lagutin, O. V. et al. *Genes Dev.* **17**, 368–379 (2003).
10. Kroll, K. L., Salic, A. N., Evans, L. M. & Kirschner, M. W. *Development* **125**, 3247–3258 (1998).
11. McGarry, T. J. & Kirschner, M. W. *Cell* **93**, 1043–1053 (1998).

Particle physics

Two is the magic number

Ken Peach

In quantum theory, the magnetic moment of the muon should be twice the value calculated classically, although in fact their ratio is not exactly two. Theory and experiment disagree on quite how far from two it is.

There are essentially three ways of exploring the strange phenomena that might occur at very high energies or, equivalently, at very high temperatures: build a powerful accelerator to produce new particles (as was done for the discoveries of the *W* and *Z* particles and the top quark); search in the present-day Universe for any 'fingerprints' of exotic particles that might have existed fleetingly in the first moments after the Big Bang; or make precise measurements of those particle properties that might be influenced by high-energy quantum processes. In this last category, one such measurement is that of the anomalous magnetic moment of the muon. Its latest value, reported by G. W. Bennett and colleagues¹, does not match the theoretical prediction as well as might be expected — and might be a hint of something as yet unknown.

The muon is, as far as we know, a truly elementary particle, similar in almost all respects to the electron except that it is some 200 times heavier. Like the electron, the muon has negative charge, and has a positively charged antiparticle. Both the electron and muon have one half-unit of the quantum property known as spin. Any spinning charge generates a magnetic field aligned

along its spin axis, so the muon and the electron have a magnetic moment, which can be calculated semi-classically from the particle's spin, charge and mass. But when the first measurements were made of the magnetic moment of the electron in the 1920s, its magnitude was found to be twice the expected value. However, one of the triumphs of Dirac's formulation of a quantum theory consistent with special relativity was that it predicted that the electron's magnetic moment would be exactly twice the semi-classical value. The ratio of the two is usually known as *g*.

In fact, life is a little more complicated. A particle's properties can be modified by other quantum effects — such as the emission and absorption of particles on short timescales, allowed by Heisenberg's uncertainty principle and illustrated in the 'loop diagram' of Fig. 1a. The magnetic moment of the electron or muon is slightly modified in this way and deviates from the expected value of two, so that $(g-2)/2$ is not exactly zero. This quantity is known as the anomalous magnetic moment and can be calculated in principle. Any significant difference between the calculated value of the anomalous magnetic moment of either the electron or muon and the corresponding

measured value is a clear indication that there must be something else, perhaps some as yet unknown particles, that should be included in the calculation (Fig. 1a).

Because the uncertainty principle allows very energetic fluctuations, provided that the timescale is suitably short, much higher energy scales can be probed through measuring the anomalous magnetic moment than can be reached directly. The effective energy scale explored is governed by two factors: the precision with which the anomalous magnetic moment can be measured, and the accuracy of the theoretical calculations. Both of these present enormous technical challenges.

On the experimental side, Bennett *et al.*¹ — the Muon ($g-2$) Collaboration, at Brookhaven National Laboratory in New York — have measured the anomalous magnetic moment of the negative muon (a_{μ^-}) with an astonishing accuracy of less than one part in a million: $a_{\mu^-} = 0.0011659214 \pm 0.0000000008$. Six months ago the same group published their result² for positive muons with the same level of accuracy and in very good agreement with previous measurements. Although this is a tricky experiment to perform, the principle behind it is fairly simple. Muons are collected inside a storage ring (Fig. 1b), constrained to move in a circular orbit by a ring of magnets. Initially, the direction of the muon spins is aligned with their velocity. But as they travel round the ring, the direction of the muon spins changes a little faster than the direction of their motion. The resulting misalignment (which grows by about 12° for each complete orbit) is directly proportional to the deviation of the g -factor from two. In fact, the muons only live a short while, decaying into an electron and two other particles called neutrinos. The direction in which the electron emerges is correlated with the direction of the decaying muon's spin, enabling measurement of the precession, and hence of a_{μ^-} (Fig. 1c).

The challenge for the theorists is equally great. There are three sets of diagrams to be incorporated in the calculation. The largest effects come from electrodynamics (processes involving photons, such as Fig. 1a); calculations have been done with up to five 'loops' in the process, to ensure that the theoretical uncertainties are smaller than the experimental errors. Weak processes (involving W and Z particles) are also significant, and calculations with up to two loops have been performed. These are computational challenges. However, the greatest uncertainty arises from loops involving other, strongly interacting particles. The strength of their interactions is such that standard computational techniques (exemplified by Fig. 1a) do not work. As more computer power becomes available, the calculation of these effects will eventually be possible. But, until then, they can only be estimated by using measurements of other processes with a similar structure.

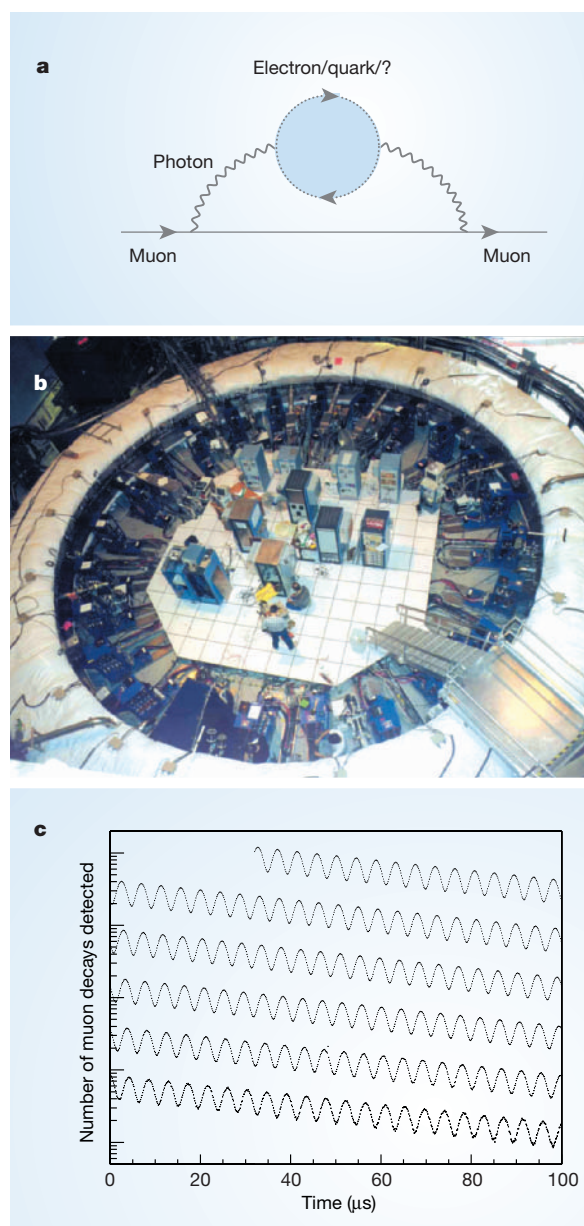


Figure 1 Loop corrections and the anomalous magnetic moment of the muon. **a**, The uncertainty principle allows momentary fluctuations such as the emission of a photon by a muon, provided that the photon is reabsorbed by the muon some short time later. Similarly, the photon itself might fluctuate into two other particles (electrons, quarks or perhaps something unknown) that then recombine. Theorists must try to estimate the effects of all possible loops on a particle's properties. **b**, This magnetic storage ring at Brookhaven National Laboratory, New York, is fed with muons. As the particles circle the ring, they gradually decay, releasing electrons that are detected by one of the 24 detectors (blue) lining the inside of the ring. **c**, The number of electrons detected varies as the muon spin direction precesses and gradually decreases as the muon population inside the ring becomes depleted. Starting from the top, the precession can be traced, in 100- μ s segments, over a period that is equivalent to about eight times the average muon lifetime. Such precision gives an accurate measurement of the anomalous magnetic moment of the muon. (The data shown are for positively charged muons².)

There have been heroic efforts to work out the contribution of pions, which are made from a quark and an anti-quark and interact strongly. One method uses data from collisions of electrons and anti-electrons (see, for example, ref. 3); another uses data from the decays of τ particles⁴ (an even heavier relation of the electron) into pions. The two theoretical calculations are not in precise agreement — they differ beyond the eighth decimal place — for reasons that are not yet understood. Nevertheless, there are indications of a discrepancy between the measurements and the theoretical expectations (at the level of 2.7 and 1.4 standard deviations, respectively). If this discrepancy is confirmed by further measurements and more refined calculations, then there is clear evidence that a new ingredient is needed in the recipe.

What might this new ingredient be? There are, of course, many theories (and many theorists) trying to answer that question.

One of the most popular ideas is that there is a new species of very massive particle, called supersymmetric particles — one consequence of their existence being a change in the anomalous magnetic moment of only a few parts in a billion. There are many other reasons why theorists would welcome such an addition to the standard model of particle physics. And there are compelling reasons to expect that, if supersymmetric particles exist, they will be found at the Large Hadron Collider, now being built at CERN in Switzerland.

Ken Peach is in the Department of Particle Physics, CLRC Rutherford Appleton Laboratory, Chilton, Didcot OX11 0QX, UK.
e-mail: k.j.peach@rl.ac.uk

1. Bennett, G. W. *et al.* Preprint at <http://arxiv.org/abs/hep-ex/0401008> (2004).
2. Bennett, G. W. *et al.* *Phys. Rev. Lett.* **89**, 101804 (2002).
3. Nyffeler, A. Preprint at <http://arxiv.org/abs/hep-ph/0305135> (2003).
4. Davier, M. *et al.* *Eur. Phys. J. C* **31**, 503–510 (2003).



100 YEARS AGO

The February number of *Knowledge and Scientific News* gives an account of what appears to be the first successful achievement of artificial flight, by Messrs. Orville and Wilbur Wright. That these brothers have been successful in gliding experiments performed under gravity is well known, but they now appear to have succeeded in raising themselves from the ground by a motor-driven machine which, after running along a mono-rail for 40 feet, rose into the air, and was driven in the face of a gale blowing at about 25 miles an hour... In the last trial the machine flew half a mile relative to the air, or 852 feet relative to the ground. It is sincerely to be hoped that this success will not, as in so many previous instances, be followed by a fatal accident.

From *Nature* 18 February 1904.

50 YEARS AGO

The recent exchange of views in *Nature* regarding the biosynthesis of proteins prompts some comments from a geneticist. Campbell and Work take note of "the two main streams of thought on protein synthesis: one derived from the study of isolated enzyme systems and suggesting a stepwise coupling of many small peptide units; the other based on the study of genetic inheritance of protein specificity and preferring synthesis on templates, each template being specific for a single protein structure and probably identifiable with a gene". While not rejecting the latter view (one gene—one protein), they point to some of the difficulties with which it is confronted, and further suggest the possibility that a synthesis of the two ideas will be found to fit the facts... The similarity of the conclusions drawn from the two aspects of this work, that with *Drosophila* and that with *Neurospora*, is notable. In connexion with the previous discussion in *Nature*, it might be suggested that the synthesis of non-specific precursor occurs in a stepwise fashion, while specificity is the function of template action. If this is the case, the gene–template relationship cannot be simple... and perhaps the template itself is synthesized by a similar two-stage process. The genetic control of protein synthesis would in any event be indirect and complex, perhaps approaching the level of complexity of the genetic control of morphogenesis.

From *Nature* 20 February 1954.

Malaria

A changed climate in Africa?

Christopher Thomas

The rise of malaria in Africa is a subject of much debate. A new analysis emphasizes the influence of rainfall, but there appear to be few areas where climate has been a major driver of this change.

After decades of decline, malaria has been on the rise in many parts of Africa — an estimate¹ by the World Health Organization is that, in some parts of the continent, malaria mortality in young children almost doubled from the 1980s to the 1990s. The disease causes some 3,000 deaths each day and imposes huge losses in economic productivity².

Is this resurgence a sign of increased transmission caused by climate change? Probably not, according to results presented by Small, Goetz and Hay³. Writing in the *Proceedings of the National Academy of Sciences*, they describe their analyses of trends in climate suitable for the regular transmission of malaria in Africa between 1911 and 1995. Their conclusion does not imply that future climate change will not affect transmission, but it does focus attention on other contemporary trends.

Several studies have projected that global climate change will increase future malaria transmission in Africa^{4–6}. However, the link between contemporary changes in malaria and climate is hotly disputed. Alternative explanations such as an increase in parasite resistance to the front-line drugs since the 1960s, poverty and a decline in many African health services are cited as more likely causes^{7–9}. Undoubtedly, a mix of such reasons is behind the rise in malaria. But identifying the prime factors will help greatly in planning control measures.

One problem dogging the interpretation of changes in local disease patterns in relation to climate is that meteorological recording stations are sparsely distributed in many parts of the continent, so that long-term records from the same locality are rare. As an alternative, researchers have used so-called 'climate surfaces' — maps interpolated from these sparse data. But these surfaces often provide only a coarse representation of climate, and their usefulness in relation to the scale of the malaria data has been open to question^{10–12}. To eliminate this mismatch, Small *et al.*³ used a malaria transmission index

calculated directly from the interpolated climate data. Refreshingly, for an area of science beset by untestable models, they then produced an 85-year 'hindcast' against which observed trends in malaria could be compared.

The malaria transmission index used by Small *et al.* was formulated in an earlier study to map zones of malaria transmission across Africa¹³. This index is based on the temperature and precipitation constraints within which the mosquito vector and malaria parasite can develop, and is pretty simple. It has nonetheless proved remarkably effective for estimating the distribution of stable



Figure 1 Water monitor — surveying mosquito breeding sites in Uganda.

ANDREW S.

Materials science

Give a shell a break

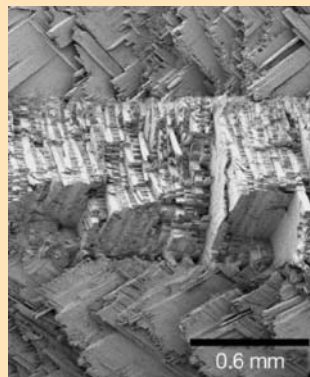
Giant conches are seldom treated with the respect they deserve. Their impressive shells are prized as holiday souvenirs, but size and aesthetics are only half the story. At the microscopic scale, they are one of nature's greatest engineering masterpieces: a stunningly intricate hierarchical architecture of inorganic crystals, interwoven with organic molecules.

In search of fresh insight into the synthesis of such biomineral composites, and clues for how to design synthetic versions, Xiao-Wei Su and colleagues have performed a series of experiments with shells of *Strombus gigas* (below), the giant queen

conch of the Caribbean. They drilled holes in the shells of live juvenile specimens, then, through X-ray diffraction and electron microscopy studies, monitored how the animals repaired their shell tissue (*Chem. Mater.* **16**, 581–593; 2004).

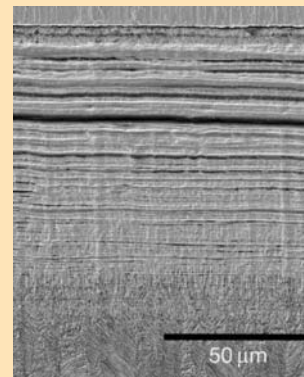
In natural shell growth, an organic outer layer (the periostracum) is deposited first. It remains unmineralized but provides a base on which the mineralized shell is deposited.

First, a micrometre-thick layer of perpendicularly oriented, elongated crystals forms, abutting the periostracum. Then comes the body of the shell,



which grows to a thickness of a few millimetres and has a three-layer 'crossed-lamellar' structure (left image, above).

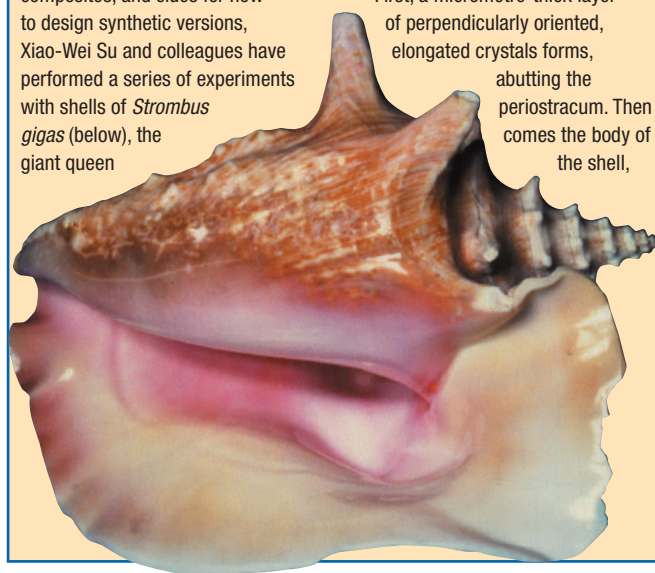
Su *et al.* found that the process of shell repair is somewhat different from natural growth. In a wounded conch, 24 hours after damage, a transparent organic membrane was deposited over the drilled hole. Once the membrane was in place, Su *et al.* saw that fine crystallites of aragonite — a particular crystal form of calcium carbonate — nucleated rapidly on the organic membrane. Each conch generated a 100- μm depth of such abnormal tissue in the damaged area (right image). After 6–8 days, elongated



crystals were deposited, oriented perpendicular to the direction of shell growth. As in normal shell development, the setting down of these elongated crystals preceded the formation of the crossed-lamellar microstructure (seen at the bottom of the image).

In uncovering the development of the unusual organic–inorganic layers in shell repair, Su *et al.* have provided a window onto the complex process of shell formation. It remains to be discovered how the interplay of organic and inorganic components is controlled at the molecular level, in conch shells as well as in other mineralized structures.

Rosamund Daw



malaria (as opposed to epidemics at unstable fringes) at coarse scales and is in operational use throughout the continent.

Using this index, Small *et al.* calculated the annual transmission suitability for each half-degree latitude–longitude grid cell (about 55 km $\times$ 55 km) across Africa from 1911 to 1995. Time-series analysis revealed that both positive and negative trends were restricted to a few limited zones, with only one (southern Mozambique) showing a consistent increase in climatic suitability for transmission. Obviously, a model at this spatial resolution will miss finer-scale patterns, for instance in climatically varied mountain areas, and will produce transmission suitability maps specific to the index used. Nonetheless, it might be expected that more widespread positive changes would have been found if climate has been a major driver of change in transmission in this period.

A notable finding, however, was that in those areas showing positive significant trends, precipitation, not temperature, drove most changes. Mathematical models of malaria transmission, often used to project

events under changed climate conditions, are based mainly on temperature-dependent processes and incorporate precipitation only as a simple global threshold sufficient for mosquito breeding⁵. Clearly this needs to be improved, but we have very little empirical understanding of how rainfall, humidity and their interactions with temperature influence vector populations. Moreover, owing to natural variability, it is difficult to identify robust signals in precipitation patterns from climate models. This is particularly so in Africa¹⁴ — a landmass that is strongly influenced by El Niño, the episodic disruption of the ocean–atmosphere system in the tropical Pacific which has a large-scale influence on weather and climate.

As far as climate change is concerned, then, the main message from Small *et al.*³ is that malaria transmission needs to be understood in terms of precipitation as well as temperature. No doubt climate models will continue to improve and we can look forward to refinement in the projection of these parameters. Regrettably, however, knowledge of the basic ecology of malaria transmission lags behind — for instance,

we cannot yet relate absolute mosquito abundance to climate⁶. The urgent need is for progress on the entomological front to guide future modelling work on transmission. ■

Christopher Thomas is at the Institute of Ecosystem Science, School of Biological and Biomedical Sciences, University of Durham, South Road, Durham DH1 2TY, UK.

e-mail: c.j.thomas@durham.ac.uk

1. World Health Organization *The African Malaria Report*, 2003 WHO/CDS/MAL/2003.1093 (WHO, Geneva, 2003).
2. Sachs, J. & Malaney, P. *Nature* **415**, 680–685 (2002).
3. Small, J., Goetz, S. J. & Hay, S. I. *Proc. Natl Acad. Sci. USA* **100**, 15341–15345 (2003).
4. Lindsay, S. W. & Martens, W. J. M. *Bull. World Health Organ.* **76**, 33–45 (1998).
5. Martens, P. *et al.* *Glob. Environ. Hum. Policy Dimensions* **9**, S89–S107 (1999).
6. Rogers, D. J. & Randolph, S. E. *Science* **289**, 1763–1766 (2000).
7. Snow, R. W., Trape, J. F. & Marsh, K. *Trends Parasitol.* **17**, 593–597 (2001).
8. Bodker, R., Kisinza, W., Malima, R., Msangeni, H. & Lindsay, S. *Glob. Change Hum. Health* **1**, 134–153 (2000).
9. Hay, S. I. *et al.* *Trends Parasitol.* **18**, 530–534 (2002).
10. Hay, S. I. *et al.* *Nature* **420**, 628 (2002).
11. Hay, S. I. *et al.* *Nature* **415**, 905–909 (2002).
12. Patz, J. A. *et al.* *Nature* **420**, 627–628 (2002).
13. Craig, M. H., Snow, R. W. & le Sueur, D. *Parasitol. Today* **15**, 105–111 (1999).
14. Hulme, M., Doherty, R., Ngara, T., New, M. & Lister, D. *Clim. Res.* **17**, 145–168 (2001).

Superconductivity

Shine a light

Michael Norman

Copper oxides superconduct at unusually high temperatures. New evidence from optical studies highlights the nature of the many-body interactions involved.

Throughout the history of superconductivity, optical spectroscopy — through the scattering of light by a material — has been a vital tool. It was the existence of a gap in the excitation-energy spectrum of electrons, first observed in optical studies, that set Bardeen on the path to the celebrated Bardeen–Cooper–Schrieffer theory of superconductivity. That theory, in which electrons move as Cooper pairs, is now the established description of the low-temperature phenomenon. On page 714 of this issue, Hwang *et al.*¹ have again demonstrated the power of optics to reveal fundamental information about superconductivity. This time, however, the revelations pertain to the underlying interactions in high-temperature superconductivity.

For copper oxides, the transition temperature at which they become superconducting is much higher than in other materials — hence the name high-temperature superconductors. Hwang *et al.* looked at the so-called optical self-energy of a bismuth-containing copper oxide (known as Bi-2212). This self-energy quantifies the deviation of the measured energy spectrum of electrons in the material from that predicted by the simple Drude theory of elementary metals. In the Drude theory, electrons are treated as hard spheres that travel in straight lines between collisions; deviations from this behaviour are thus a measure of the strength of the many-body interactions that the electrons undergo.

Hwang *et al.*¹ find that, for temperatures above the superconducting transition temperature, this self-energy has a rather broad and featureless distribution, up to very high electron excitation energies. Such behaviour is inconsistent with that expected if the moving electrons are interacting with phonons — vibrations in the atomic lattice — as occurs in low-temperature superconductors. Instead, Hwang *et al.* conclude that the self-energy distribution is due to interactions between the electrons themselves. This finding supports a large body of theoretical work that argues that copper oxide superconductors are indeed fundamentally different from their low-temperature counterparts. As Hwang *et al.* also suggest, this optical self-energy can be considered as a measure of the ‘glue’ that binds electrons into Cooper pairs, binding that in turn gives rise to superconductivity.

Even more interesting are their results for

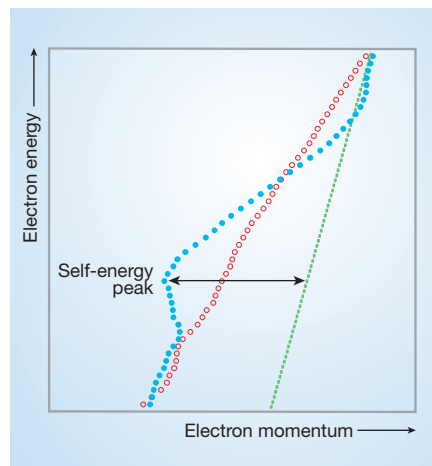


Figure 1 Self-energy and the superconductor. Measured in photoemission experiments² on optimally doped Bi-2212, the dispersion relation between the energy and momentum of electrons in the material is shown at different temperatures. The transition temperature at which the material becomes superconducting is 90 K; data are shown for temperatures each side of that mark, at 40 K (blue dots) and 140 K (red circles). The green dashed line is the bare theoretical estimate. At 40 K, an S-shaped curve appears in the dispersion; the maximum displacement of the data curve from the theory corresponds to a peak in the electron self-energy. This effect is analogous to that now observed by Hwang *et al.*¹ in optical studies. These authors argue¹ that the peak is caused by the interaction of electrons with a magnetic resonance.

Bi-2212 at lower temperatures. As the material is cooled, a sharp peak develops in the optical self-energy. This behaviour is analogous to that seen for copper oxides in other experiments based on photoemission. In those experiments, the ‘dispersion relation’ between the energy and momentum of the electrons is measured at different temperatures. Figure 1 shows the typical variation² of energy with momentum for a copper oxide sample (optimally doped) at two different temperatures: at 140 K, which is above the transition temperature for the sample, and at 40 K, which is below the transition temperature. The deviation of the measured dispersion from its bare predicted value is proportional to the electron self-energy. Above the transition temperature, the dispersion is rather featureless, although it is shifted relative to its bare value; below the transition temperature, a characteristic

S-shaped curve develops. The maximum amplitude of this ‘S’ represents the peak in the electron self-energy.

Such S-shaped dispersions are well known in many-body physics. They are a consequence of the interaction of the electrons with a collective mode that is sharp in energy. Although Lanzara *et al.*³ have advocated that this mode is a phonon, it has been suggested by other groups^{2,4} (including mine) that the mode is a magnetic resonance, evidence of which has been seen in inelastic neutron-scattering experiments⁵.

The data put forward by Hwang *et al.*¹ strongly support the magnetic-resonance interpretation of the self-energy peak. They have tracked the optical self-energy peak at various temperatures over a wide range of chemical doping (that is, varying the amount of mobile charge-carriers added to the sample, which affects its performance as a superconductor) and they find that the peak is directly correlated with the magnetic resonance. They also find that, as the doping increases, the peak eventually disappears at a doping level for which the superconductivity is still strong — electrons are still joined into Cooper pairs, so the peak itself, they suggest, cannot represent their binding glue.

Although this assertion is true, it should be noted that the magnetic resonance is a feedback effect associated with the onset of superconductivity — thus it does not cause superconductivity but is caused by it⁶. The reason for its disappearance at high doping levels is the expectation that, as the many-body interactions weaken with increased doping (which is supported by the optics¹ and photoemission⁴ measurements), the resonance will in turn weaken and disappear.

But the fact that the optical self-energy peak grows continuously out of the broad background as the temperature is lowered, and that it has, moreover, a magnitude comparable to that of the background, indicates that the background ‘glue’ and the peak have a common origin. Because the peak is magnetic in nature, this implies that the glue is magnetic as well. Consequently, Hwang *et al.*¹ have provided strong evidence that the origin of high-temperature superconductivity in copper oxides is indeed connected with magnetic correlations. This work will be a great encouragement to those pursuing a magnetic pairing model for copper oxide superconductivity. ■

Michael Norman is at the Materials Science Division, Argonne National Laboratory, Argonne, Illinois 60439, USA.

e-mail: norman@anl.gov

1. Hwang, J., Timusk, T. & Gu, G. D. *Nature* **427**, 714–717 (2004).

2. Kaminski, A. *et al.* *Phys. Rev. Lett.* **86**, 1070–1073 (2001).

3. Lanzara, A. *et al.* *Nature* **412**, 510–514 (2001).

4. Johnson, P. D. *et al.* *Phys. Rev. Lett.* **87**, 177007 (2001).

5. Rossat-Mignod, J. *et al.* *Physica C* **185–189**, 86–92 (1991).

6. Abanov, A. & Chubukov, A. V. *Phys. Rev. Lett.* **83**, 1652–1655 (1999).

Evolution

Adapting to life on campus

Evolution **58**, 166–174 (2004)

When colonizing a new habitat, a species can encounter new selection pressures. A population of dark-eyed juncos (*Junco hyemalis*) that left its mountain habitat and set up home at the University of California, San Diego, has allowed researchers to see this effect in action. For the past two decades, the birds' sexual plumage displays have been evolving at a surprising rate.

Male juncos signal their quality to mates through the amount of white plumage in their tail. But the San Diego birds have 22% less white in their tail than do their mountain counterparts, reports Pamela J. Yeh. As the birds arrived in 1983, this is impressively rapid evolution.

The move to a warmer climate, with a longer breeding season, may have forced males to concentrate on proving their worth as fathers, rather than as Adonises, Yeh suggests. With the opportunity for several clutches each summer, females may favour males who helped to raise previous broods that they had sired. Coupled with the campus population's low density, this reduces the need for striking plumage and may have driven its decline.

This could not have occurred without such selection, says Yeh, and the campus population is too big for such a large genetic change to have occurred randomly.

Michael Hopkin

Chemistry

On with the flow

J. Am. Chem. Soc. **126**, 1569–1576 (2004)

Water can form a cage-like lattice that traps other molecules, creating solid 'clathrate hydrates'. For example, methane hydrates, trapped beneath the oceans, are believed to be the most abundant form of natural methane. But hydrates are bad news when they build up in pipelines and block gas flow. Although chemical additives can inhibit clathrate formation, they stop working if the pipes get cold enough for long enough.

Through a computational screening process, Mark T. Storr *et al.* have designed a new class of clathrate inhibitor. An exemplar, tributylammonium propylsulphonate (TBAPS), performed at least as well as one of the most commonly used inhibitors in a series of tests.

According to Storr and colleagues' analysis, one end of the TBAPS molecule attracts water molecules, which weakens the surrounding lattice; the opposite end repels water, strengthening the lattice in that region. This distortion effectively delays the formation of clathrate crystals. The authors

believe that this new compound can be developed further to substantially increase its activity.

Mark Peplow

Transcription

It's about time

Proc. Natl Acad. Sci. USA **101**, 1200–1205 (2004)

At a sufficiently high cell density, yeast cultures synchronize their efforts during the so-called respiratory cycle: in 40-minute bouts within the cell cycle, they alternate respiratory tasks, which generate energy, with reductive functions, during which they produce cellular constituents. While hunting for genes that support these oscillations, Robert R. Klevecz and colleagues uncovered genome-wide rhythms in gene transcription.

The researchers scrutinized gene activity in yeast by using gene chips, and found three distinct intervals in each respiratory cycle when most gene transcription occurred. The vast majority of genes were expressed during one of two periods in the reductive phase of the cycle, whereas the rest were active during a single period in respiration. Even those genes previously thought to be important in cellular housekeeping, and 'constitutively' expressed, showed small oscillations.

Klevecz *et al.* further showed that, every 40 minutes, a burst of DNA replication followed the respiratory phase in a constant fraction of the yeast culture. The authors believe that restricting DNA synthesis to the reductive phase of the cycle, when potentially harmful reactive-oxygen species are less abundant, may have evolved to reduce oxidative damage to DNA. Moreover, they propose that multiples of 40-minute oscillations underlie cell division.

Marie-Thérèse Heemels

Immunology

Toxic shock in the dock

Cell **116**, 367–379 (2004)

Toxic shock syndrome, caused by the bacterium *Streptococcus pyogenes*, is one of the severest forms of septic insult. Normally, the bacterium brings on a mild throat or skin infection. But occasionally it goes on to cause blood-vessel leakage and organ failure, with rapidly fatal results. Working with both *in vitro* and *in vivo* models, Heiko Herwald *et al.* have characterized the chain of events by which *S. pyogenes* wreaks havoc in its unfortunate host.

The researchers find that a protein released from the surface of the bacterium, 'M protein', attaches itself to fibrinogen, a component of blood. The complex of M protein and fibrinogen activates neutrophils, a type of white blood cell, which then release heparin-binding protein — the molecule that produces the ultimately catastrophic failure of blood vessels.

Herwald *et al.* went further, and by pinpointing a peptide that interferes with the fibrinogen–neutrophil interaction, they identified where the chain might be broken. As they say, 30–70% of patients die despite intensive efforts to save them, and more effective treatments for the syndrome are badly needed.

Laura Nelson



Materials science

Snow has no weakest link

Phys. Rev. E **69**, 011306 (2004)

Skiers and mountain guides know that snow is treacherous stuff, but H. O. K. Kirchner and colleagues have demonstrated just how treacherous. One never knows, they say, when snow is going to crack — the process that typically triggers an avalanche.

The researchers quantify this unreliability of snow by measuring its so-called Weibull modulus, a measure of the variability of a material's strength. The smaller this parameter is, the broader the spread of different fracture strengths for ostensibly identical samples. High-performance steel has a modulus of 22, and for normal ice the value is 3–10. For snow, however, it is 1.5 ± 0.5 . This means that even very low stresses have a chance of fracturing snow.

On the other hand, Weibull's original analysis postulated that the strength of a brittle material is limited by the strength of the weakest link in a random distribution of flaws. This implies that the strength declines as the sample gets larger (increasing the chance of it containing a weak link). But Kirchner *et al.* find that snow has no such size dependence — which means that lab-scale measurements of its mechanical properties can probably be extrapolated to the much larger scales relevant to avalanche prediction.

Philip Ball

Cancer cells compress intratumour vessels

Pressure from proliferating cells impedes transport of therapeutic drugs into tumours.

The delivery of therapeutic drugs to solid tumours may be impaired by structural and functional abnormalities in blood and lymphatic vessels¹. Here we provide evidence that proliferating cancer cells cause intratumour vessels to compress and collapse. By reducing this compressive mechanical force and opening vessels, cytotoxic cancer treatments have the potential to increase blood perfusion, thereby improving drug delivery.

Elevated interstitial fluid pressure — a hallmark of solid tumours — is commonly assumed to compress intratumour vessels. However, the interstitial fluid pressure is about equal to the microvascular pressure in tumours, making it unlikely that the collapse of permeable vessels is mediated by fluid pressure². We therefore investigated whether proliferating cancer cells might cause compression of tumour vessels and, if so, whether relieving this compression would allow them to open and regain their functionality.

To relieve compression, we targeted tumour cells with diphtheria toxin, which is much less cytotoxic to mouse than to human cells³ (see supplementary information for methods). Diphtheria toxin had no effect on murine tumours (T-241 type), normal murine tissue (liver) or endothelium, but caused apoptosis and time-dependent regression in human-tumour xenografts (HSTS-26T type) grown in mice (results not shown).

We found a greater fraction of blood and lymphatic vessels with an open lumen in diphtheria-toxin-treated human tumours than in controls treated with PBS buffer (Fig. 1a–c). Treatment with diphtheria toxin also causes vessels to have a rounder cross-section in the human tumours but has no effect on vessel morphology in the murine tumours. Compared with collapsed vessels, open vessels are surrounded by fewer cells, suggesting that the local compressive force is lower (Fig. 1d). The diameter of perfused blood vessels increased in the human tumours after treatment with diphtheria toxin and, in contrast to results obtained after treatment with taxane⁴, a greater fraction of blood vessels were perfused.

These changes in vascular morphology cannot be explained by a difference in collagen I content or in the molecular determinants of blood- and lymphatic-vessel formation and remodelling (namely, vascular endothelial growth factor, vascular endothelial growth factor-C, vascular endothelial growth factor receptor-2, angiopoietin-1 or angiopoietin-2). Vessels associated with smooth-muscle α -actin-expressing cells or

with a prominent collagen IV basement membrane are more likely to have an open lumen. Lymphatic vessels are rarely adjacent to smooth-muscle α -actin-expressing cells.

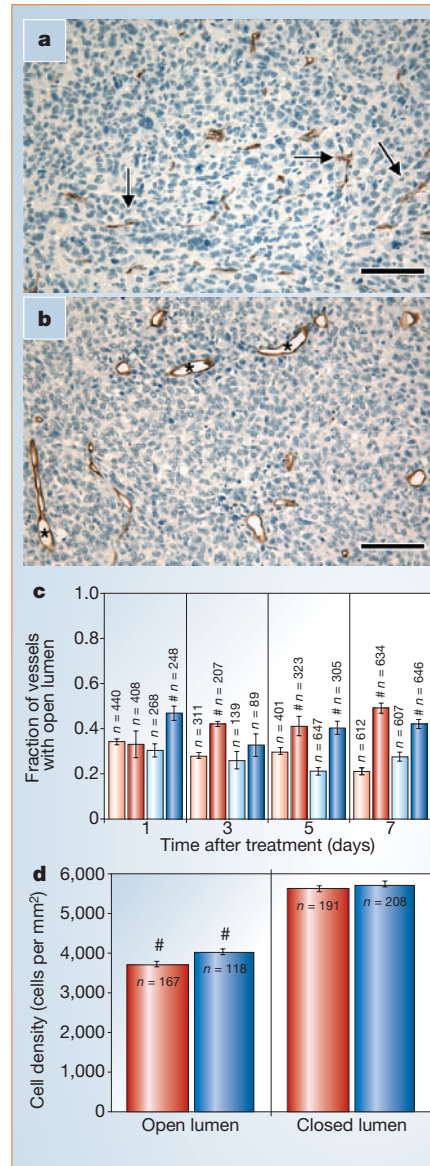


Figure 1 Collapsed blood and lymphatic vessels open after relieving compressive forces generated by cancer cells in human tumours. **a**, Blood vessels in HSTS-26T tumours remain collapsed (arrows) 5 days after treatment with PBS buffer. **b**, A greater fraction of blood vessels have an open lumen (asterisks) 5 days after treatment with diphtheria toxin. Scale bars, 50 μ m. Results were similar for lymphatic vessels. **c**, Fraction of blood vessels (red bars) and lymphatic vessels (blue bars) with an open lumen at different times after treatment with diphtheria toxin (dark bars) or PBS buffer (light bars). **d**, Blood vessels (red bars) and lymphatic vessels (blue bars) with an open lumen have a lower surrounding cell density compared with vessels that have a collapsed lumen. Hash symbols indicate that $P < 0.05$.

We next tested whether intratumour lymphatic vessels, which are ordinarily non-functional or absent^{5,6}, become functional upon opening. We found, surprisingly, that those inside diphtheria-toxin-treated tumours did not become functional, as judged by fluorescence or ferritin micro-lymphangiography⁶. It is possible that tumours permanently damage lymphatic-vessel structures, such as lymphatic endothelial microvalves, or that the lack of pulsatile blood flow in tumours inhibits lymph formation⁷.

The tumour margin, in which mechanical stress is predicted to be lower (see supplementary information), contained functional lymphatic vessels, as well as a greater fraction of open lymphatics than was found in intratumour regions (data not shown). Regions subject to lower predicted compressive forces thus contain open, functional lymphatics, whereas regions exposed to greater compressive forces contain collapsed, non-functional lymphatic vessels.

The role of a tumour's mechanical micro-environment in its progression and treatment is starting to emerge. For example, compressive forces inhibit tumour cell growth⁸ and upregulate adhesion molecules⁹. Clinically, tumours can compress large vessels and the spinal cord.

Our findings indicate that proliferating tumour cells can also cause intratumour vessels to collapse, particularly those without supportive stromal structures, leading to impaired function. Tumour-specific cytotoxic therapy may result in more efficient drug (or nutrient) delivery by decompressing collapsed vessels¹⁰. Whether these newly perfused vessels also provide additional routes for metastasis is not yet known.

Timothy P. Padera, Brian R. Stoll, Jessica B. Tooredman, Diane Capen, Emmanuelle di Tomaso, Rakesh K. Jain

Edwin L. Steele Laboratory, Department of Radiation Oncology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA

e-mail: jain@steele.mgh.harvard.edu

- Jain, R. K. *Nature Med.* **9**, 685–693 (2003).
- Boucher, Y. & Jain, R. K. *Cancer Res.* **52**, 5110–5114 (1992).
- Arbiser, J. L. et al. *Am. J. Pathol.* **155**, 723–729 (1999).
- Griffon-Etienne, G., Boucher, Y., Brekken, C., Suit, H. D. & Jain, R. K. *Cancer Res.* **59**, 3776–3782 (1999).
- Alitalo, K. & Carmeliet, P. *Cancer Cell* **1**, 219–227 (2002).
- Padera, T. P. et al. *Science* **296**, 1883–1886 (2002).
- Schmid-Schönbein, G. W. *Lymph. Res. Biol.* **1**, 25–31 (2003).
- Helmlinger, G., Netti, P. A., Lichtenbeld, H. C., Melder, R. J. & Jain, R. K. *Nature Biotechnol.* **15**, 778–783 (1997).
- Koike, C. et al. *Br. J. Cancer* **86**, 947–953 (2002).
- Reinhold, H. S. *Eur. J. Cancer* **7**, 273–280 (1971).

Supplementary information accompanies this communication on Nature's website.

Competing financial interests: declared none.

The stability of forest biodiversity

The unified neutral theory of biodiversity and biogeography¹ provides a dynamic null hypothesis for the assembly of natural communities. It is also useful for understanding the influence of speciation, extinction, dispersal and ecological drift on patterns of relative species abundance, species–area relationships and phylogeny. Clark and McLachlan² argue that neutral drift is inconsistent with the palaeorecord of stability in fossil pollen assemblages of the Holocene forests of southern Canada. We show here that their analysis is based on a partial misunderstanding of neutral theory and that their data alone cannot unambiguously test its validity.

Neutral theory in ecology¹ builds on the theory of island biogeography³, which asserts that an island or a local community approaches steady-state species richness where there is an equilibrium between the immigration of species from the much larger metacommunity source area and the local extinction of species. The dynamics of a local population are governed by birth, death and immigration events in both neutral and non-neutral models. Under neutrality, at large spatial and temporal scales, Fisher's log-series distribution is the expected steady-state distribution of relative species abundance at the speciation–extinction equilibrium in the metacommunity when the per capita birth and death rates are density independent and the same for all species, and speciation is introduced⁴.

The characteristic timescale for speciation and evolutionary changes in species is much longer than the 200-generation period studied by Clark and McLachlan. During this relatively short time interval, the large-scale metacommunity distribution of relative species abundance, particularly of common and widespread species, will change slowly or remain nearly constant¹.

Clark and McLachlan's analysis² of the fossil pollen record of the eight local tree communities assumes that the sites are undergoing drift in complete dynamic isolation. However, in common with the theory of island biogeography³, the neutral theory^{1,4} takes into account the fact that the species composition and the dynamics of local communities will be influenced by immigration from the surrounding metacommunity. When the local community is coupled dynamically to the metacommunity, local communities will covary positively in species composition, and exhibit potentially greater compositional stability than if they drifted in complete isolation, as in Clark and McLachlan's simulations.

Dispersal tends to stabilize local species composition and relative abundances because communities become dynamically coupled to one another. Immigrants belonging to the more abundant metacommunity species will arrive more frequently in the local community than immigrants of rarer species, stabilizing local assemblages and making them more similar. Very fast rates of long-distance dispersal may characterize the northward movements of several tree species following the retreat of the continental ice sheets during the Holocene of Europe⁵ and North America⁶.

One question is how stable the metacommunity surrounding these local communities has been over the past 10,000 years (the length of the fossil record considered by Clark and McLachlan²). Large changes were occurring over this time in the post-glacial forests of eastern North America. Whether or not the metacommunity provided a relatively stable backdrop for the dynamics within the sampled local communities is unclear. However, even if the metacommunity was not perfectly constant or was changing gradually and directionally, abundant metacommunity species would be expected to be major components of virtually all local communities, even communities weakly coupled to the metacommunity with relatively slow or infrequent immigration¹. A stronger argument for stability inconsistent with neutrality would be to show that local communities persisted in their compositions, despite large changes in the abundances of common species in the metacommunity in which they were embedded.

Figure 1 of Clark and McLachlan² is based on simulations of neutral dynamics using a lottery model, which the authors claim measures the divergence among sites, over time, that is expected under neutrality and which should continue to increase until diversity is lost through extinction. But we dispute this. In their model, which does not allow for extinction, the variance actually levels off with time, reaching an equilibrium. Thus, they overestimate the accumulation of among-site variance under neutrality.

Moreover, the timescale for equilibration in Clark and McLachlan's simulations is very different from the evolutionary timescales associated with equilibration of biodiversity in the metacommunity. Biodiversity is maintained in equilibrium in neutral theory because of the balance between speciation and extinction of species on very large spatial and temporal scales. The simplified lottery model used as a neutral benchmark by Clark and McLachlan² does not capture a fundamental result of unified neutral theory¹, namely that the fates of all species in the local community differ according to their unequal abundances in the metacommunity.

Clark and McLachlan note that the distributions of many plant species are strongly correlated with climate, hydrology and soil, but such evidence for niche assembly need not preclude a role for demographic stochasticity (ecological drift). But how strong are the deterministic processes structuring ecological communities relative to drift? Although the hypothesis of complete neutrality may often fail (it is, after all, a null hypothesis), we do not agree with Clark and McLachlan's assertion that the strong evidence for stabilizing forces in the palaeorecord² is sufficient to reject ecological drift. Because their lottery model prevents extinction and ignores dispersal and the dynamic coupling between local communities and the metacommunity, it cannot be regarded as a definitive test of current neutral theory.

Igor Volkov*, Jayanth R. Banavar*, Amos Maritan†, Stephen P. Hubbell‡

*Department of Physics, 104 Davey Laboratory, Pennsylvania State University, University Park, Pennsylvania 16802, USA

e-mail: banavar@psu.edu

†Dipartimento di Fisica 'G. Galilei', Università di Padova and INFN 35131 Padova, and Abdus Salam International Center for Theoretical Physics, 34014 Trieste, Italy

‡Department of Plant Biology, University of Georgia, Athens, Georgia 30602, USA and Smithsonian Tropical Research Institute, Box 2072, Balboa, Panama

1. Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton Univ. Press, Princeton, NJ, 2001).
2. Clark, J. S. & McLachlan, J. S. *Nature* **423**, 635–638 (2003).
3. MacArthur, R. H. & Wilson, E. O. *The Theory of Island Biogeography* (Princeton Univ. Press, Princeton, NJ, 1967).
4. Volkov, I., Banavar, J. R., Hubbell, S. P. & Maritan, A. *Nature* **424**, 1035–1037 (2003).
5. Petit, R. J. *et al.* *Forest Ecol. Management* **156**, 49–94 (2002).
6. Clark, J. S. *et al.* *Bioscience* **48**, 13–24 (1998).

Clark and McLachlan reply — The neutral model says that the relative abundance of a species is as likely to increase as it is to decrease, because species are ecologically identical¹. This hypothesis can be rejected if variance does not increase over time². We used a more powerful test, based on comparisons among locations, to show that variability stabilizes and, for most species, decreases over thousands of years³. The neutral model also predicts that after a perturbation, relative abundance is as likely to increase as it is to decrease.

Eastern hemlock, a species decimated throughout its range in the mid-Holocene, subsequently increased throughout eastern North America. It did not return to the same densities; it was not expected to because Holocene climates have varied. But the increase from low abundance throughout its range is inconsistent with neutral drift.

Volkov *et al.* object to our analysis, saying that dispersal causes “immigrants belonging to the more abundant metacommunity species to arrive in the local community

more frequently than immigrants of rarer species, stabilizing local assemblages". Their argument says that local populations are maintained by propagule pressure. An abundant species might increase owing to overwhelming seed supply, despite having no competitive advantage. This explanation requires abundant seed emanating from large residual populations. Of course, hemlock was not eradicated from North America. Yet there is no indication of a more abundant metacommunity that could have caused hemlock seeds to arrive more frequently in the local community than did other species. Other species were not rare — the region remained forested throughout. If there was an as yet undiscovered large hemlock population at the time, dispersal data do not support the view that the recovery came from hemlock seed that was sufficiently abundant to overwhelm other species throughout eastern North America. Only beneath hemlock crowns does its seed arrive more frequently than that of other species^{4,5}.

One explanation that is consistent with our results, field studies and models is that hemlock competes well in shaded understoreys. An increase from low density is expected in view of its shade tolerance. Differences among species contribute to patterns of diversity across gradients, during succession and with climate change. The predictability of succession, based in large part on shade tolerance, is an example of non-neutral dynamics.

Empirical data and models do not support the view of Volkov *et al.* that dispersal causes tree populations to covary over broad regions (ours exceeds 10⁵ km²). Seed produced in distant populations makes a small contribution to local density. For this reason, spatial population models predict correlation lengths close to mean dispersal distances (tens of metres)^{6,7} and inconsequential representation of lineages that might derive from distant dispersal^{8,9}. The latter prediction is consistent with genetic haplotypes¹⁰. The tight geographic coupling described by Volkov *et al.* could indeed be an outcome of Hubbell's non-spatial model, which assumes global migration of recruits that are immediately everywhere. Unlike dispersal, this assumption is not spatial.

Volkov *et al.* believe that a period of several hundred generations is not long enough to test the unified neutral theory. We stressed that our analysis is not a test of whether speciation offsets extinction; we only tested neutral dynamics³.

We agree that our model is not strictly neutral³. The lottery example was not used to 'test' for neutrality. We did not dwell on asymptotic behaviour of the model (eventually involving extinction), because it is not relevant for this example — these species co-exist. Species could be assigned identical parameters in any model. The

model simply illustrates how variances increase for identical species, a model-independent prediction. Our test was based solely on observed data.

James S. Clark, Jason S. McLachlan

Center on Global Change, Biology, and Nicholas

School of the Environment, Duke University,

Durham, North Carolina 27708, USA

e-mail: jimclark@duke.edu

1. Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton University Press, Princeton, NJ, 2001).
2. Murdoch, W. W. *Ecology* **75**, 271–287 (1994).
3. Clark, J. S. & McLachlan, J. S. *Nature* **423**, 635–638 (2003).
4. Ribbens, E., Silander, J. A. & Pacala, S. W. *Ecology* **75**, 1794–1806 (1994).
5. Clark, J. S., Macklin, E. & Wood, L. *Ecol. Monogr.* **68**, 213–235 (1998).
6. Pacala, S. W. *et al. Ecol. Monogr.* **66**, 1–43 (1996).
7. Keeling, M. J., Mezic, I., Hendry, R. J., McGlade, J. & Rand, D. A. *Phil. Trans. R. Soc. Lond. B* **352**, 1589–1601 (1997).
8. Ibrahim, K. M., Nichols, R. A. & Hewitt, G. M. *Heredity* **77**, 282–291 (1996).
9. Le Corre, V., Machon, N., Petit, R. J. & Kremer, A. *Genet. Res.* **69**, 117–125 (1997).
10. Petit, R. J. *et al. Proc. Natl Acad. Sci. USA* **94**, 9996–10001 (1997).

COMMUNICATIONS ARISING

Ecology

Living in synchrony on Greenland coasts?

Theory indicates that correlated weather may synchronize populations¹, but the extent to which this holds for non-identical, nonlinear systems is uncertain. Post and Forchhammer² claim to have shown climate-induced synchrony for musk oxen and caribou that are separated by the Greenland ice sheet. However, logical and mathematical errors undermine their finding. Whether or not large-scale weather can be a major synchronizing factor across species remains an open question.

Some caribou–musk-ox population pairs show correlated abundances, and Post

and Forchhammer ascribe this to a high 'North Atlantic Oscillation (NAO) effect ratio', $\omega_{\text{musk ox}}^{(j)}/\omega_{\text{caribou}}^{(j)}$ where $\omega_{\text{musk ox}}^{(j)}$ and $\omega_{\text{caribou}}^{(j)}$ are regression coefficients quantifying how the NAO (ref. 3) affected the growth of each population⁴. However, the effect ratio increases if $\omega_{\text{caribou}}^{(j)}$ decreases, so the authors' conclusion illogically implies that the less the climate affects caribou, the more climate-synchronized they will be with the musk oxen. Instead, it would be expected that strong climate effects on both populations were a necessary (although not sufficient) requirement for climate-induced synchrony between non-identical systems where local, uncorrelated noise dilutes the effect of shared climatic forcing.

Post and Forchhammer argue that cold winters synchronize population dynamics, reporting a correlation between the NAO and the 'yearly synchrony' (a measure of variability in abundance among populations) within a species. However, synchrony (that is, togetherness in time) requires a temporal aspect, comparing either the timing of events or rates of change, both of which require the availability of more than one time point. The authors' attempt to deduce synchrony from a snapshot of abundances at a single point in time, which is not possible.

Post and Forchhammer's chosen measure of population-size variability is mathematically not well defined, so the results can be turned either way simply by a change of measurement unit (Fig. 1). They calculated the ratio of the mean to the standard deviation of log-transformed abundances, $\text{sync}(N) = [\text{mean}_{\text{all } i}(\ln N_i)/\text{stdev}_{\text{all } i}(\ln N_i)]$, for a group of populations in a given year.

However, the mean and standard deviation cannot be compared when using log-transformed data, because the mean of the transformed values depends on the measurement unit used for the raw data, whereas

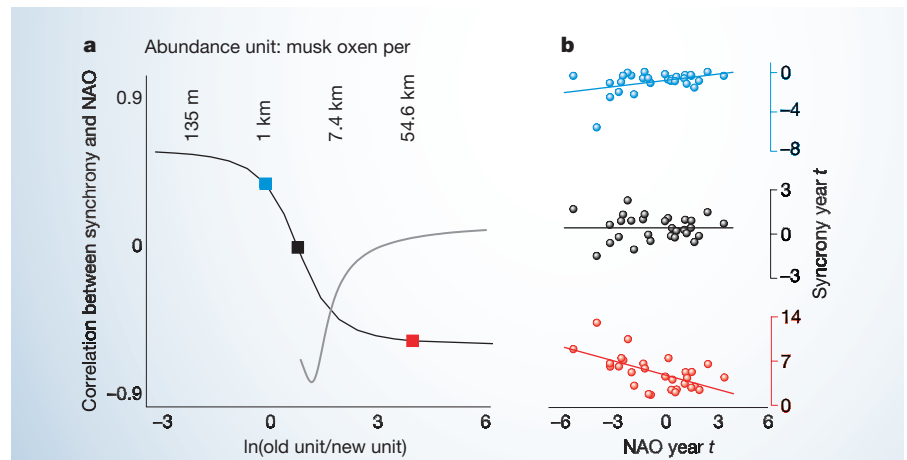


Figure 1 Post and Forchhammer's results can be turned either way by a change of measurement unit. **a**, Effect of the choice of unit of measurement on the correlation between NAO and 'yearly synchrony' of musk-ox populations. Squares correspond to scatterplots in **b** (blue, based on raw data; red, Post and Forchhammer's calculations; black, a measurement unit that eliminates the correlation altogether). Post and Forchhammer's adjustment of the unit so that all log-values are positive does not make their synchrony measure well defined, as highlighted by a subset of the data in grey. **b**, Scatterplots against NAO of 'yearly synchrony', using units indicated by the squares in **a**.

the standard deviation does not. This is because changing the units corresponds to multiplying the data with a constant (a), which adds a constant to the log values: $\ln aN_i = \ln a + \ln N_i$. The resulting change in 'synchrony' depends both on $\text{stdev}_{\text{all } i}(\ln N_i)$, which is specific to each year's set of abundances, and on the choice of unit (specifically, $\text{sync}(aN) = \text{sync}(N) + \ln(a)/\text{stdev}_{\text{all } i}(\ln N_i)$). Therefore, a change of unit may qualitatively reverse comparisons of synchrony between data sets.

This problem is not resolved by addition of a constant to all log-transformed musk-ox abundances to ensure that these are positive before calculating 'synchrony' (Post and Forchhammer, personal communication), as the relationship between mean $\ln aN$ and $\text{stdev} \ln aN$ remains entirely arbitrary. Biological conclusions should not be affected by whether American, metric or other units are used.

Post and Forchhammer's analyses show an apparent tendency for cross-species correlation to decrease with increasing interpopulation distance. However, the few strong correlations describe concurrent trends over decades, rather than the year-to-year variation that was Post and Forchhammer's focus. Plotting cross-species correlation against each ($\omega_{\text{musk ox}}^{(i)}$, $\omega_{\text{caribou}}^{(j)}$) pair shows no consistent pattern. We also correlated growth rates instead of raw abundances and found unsystematic and weak correlations. Although other approaches might be more successful, the data may not be sufficiently precise or relevant to detect the phenomenon if it exists. We conclude that there is currently no proper evidence of climate-induced synchrony between musk oxen and caribou on Greenland.

Jon Olav Vik*, Nils Chr. Stenseth*, Giacomo Tavecchia*†, Atle Mysterud*, Ole Chr. Lingjærde†

*Centre for Ecological and Evolutionary Synthesis, Department of Biology, and †Department of Informatics, University of Oslo, PO Box 1050 Blindern, 0316 Oslo, Norway
e-mail: n.c.stenseth@bio.uio.no

†Present address: IMEDA-UIB/CSIC-c Miquel Marques 21, 07190 Esporles, Spain

1. Moran, P. A. P. *Aust. J. Zool.* **1**, 291–298 (1953).
2. Post, E. & Forchhammer, M. C. *Nature* **420**, 168–171 (2002).
3. Hurrell, J. W. *Science* **269**, 676–679 (1995).
4. Forchhammer, M. C., Post, E., Stenseth, N. C. & Boertmann, D. M. *Population Ecol.* **44**, 113–120 (2002).

Post and Forchhammer reply — Vik *et al.* question whether we documented spatial synchrony between caribou and musk oxen from Greenland, and whether spatial synchrony within each species related to the North Atlantic Oscillation (NAO)¹. Attributing spatial synchrony to climate is difficult but possible², and the questions raised by Vik *et al.* are readily addressed. Contrary to their incorrect statement of our

definition of the NAO effect ratio¹, a strong climatic effect on any pair of populations is not a requisite of climate-induced synchrony. As Moran³ argued, and as our analysis illustrated¹, populations may be synchronized if climate influences each of them similarly, regardless of the magnitude of that influence. Moreover, the standardized NAO effect ratio is associated statistically with the degree of climatic correlation across populations⁴ and hence the degree of synchrony between populations⁵.

As stated previously¹, we used cross-population covariance (CV) to produce a time-series index of spatial synchrony, an approach validated in empirical⁵ and theoretical⁶ studies, which have demonstrated the relationship of CV to population synchrony⁷. The simplest test of whether the use of log-transformed data confounds our results is to compare them with results obtained using raw (not log-transformed) data. The correlation between the NAO and 1/CV of the raw musk-ox data⁸ ($r = -0.57$, $P = 0.001$) matches exactly the correlation between the NAO and 1/CV of the log-transformed musk-ox data ($r = -0.59$, $P = 0.001$). Similarly, results do not vary for caribou, using log-transformed ($r = 0.35$, $P = 0.002$) or raw ($r = 0.24$, $P = 0.04$) data. Hence, log-transformation does not influence relationships between the NAO and spatial synchrony.

Moreover, our results were not influenced by addition of a constant to the log-transformed musk-ox data, which Vik *et al.* describe as analogous to changing units of abundance. Such a problem would be apparent if the means of the N_i , $\ln(N_i)$, or $[\ln(a) + \ln(N_i)]$ showed significant and inconsistent correlations with the NAO, but none did (r_i values of 0.14, 0.07 and 0.07, respectively; all P values ≥ 0.50). Vik *et al.* obtained different results because their direct log-transformation of the decimal-form musk-ox data produced negative values, giving statistically invalid CVs^{9,10}. We added 4 to the log-transformed musk-ox data to convert negative values to positives before calculating the CV precisely to avoid a spurious correlation.

If the NAO–spatial synchrony correlation were, in fact, influenced by the use of log-transformed data, such an artefact should be apparent in two ways, neither of which is discussed by Vik *et al.* First, the sign of the correlation between the NAO and musk-ox spatial synchrony should change with addition of constants greater than 4. Second, adding constants to log-transformed caribou data (to which none was originally added¹) should also alter the NAO–caribou synchrony correlation. We checked this by adding constants of up to 10, and in neither case was the correlation altered (Fig. 1).

We conclude that our previous results still stand and that Vik *et al.* cannot offer a

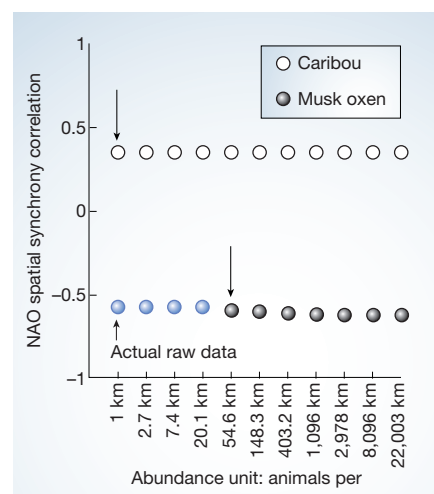


Figure 1 Addition of constants from zero to ten (corresponding to abundance units shown on the x-axis) to $\ln$ -transformed density estimates does not influence the sign of the correlation between the NAO index and spatial synchrony for musk oxen or caribou. Points designated by an arrow are the correlations reported in our original analysis¹. For the $\ln$ -transformed musk-ox data, addition of constants less than four, as done by Vik *et al.*, produces spurious correlations because the resulting negative values do not give statistically valid coefficients of variation^{9,10}. Thus, the blue points indicate the correlations obtained when using the raw musk-ox data⁸ with constants of zero to three added; note that the point denoted as raw data in Fig. 1 of Vik *et al.* is in fact a direct $\ln$ -transformation of the actual raw data.

means of analysing or an alternative explanation for spatial synchrony within and across these species.

Eric Post*, Mads C. Forchhammer†

*Department of Biology, The Pennsylvania State University, 208 Mueller Lab, University Park, Pennsylvania 16802, USA

e-mail: esp10@psu.edu

†Department of Population Biology, Biological Institute, University of Copenhagen, 2100 Copenhagen, Denmark

1. Post, E. & Forchhammer, M. C. *Nature* **420**, 168–171 (2002).
2. Cattadori, I. M., Merler, S., & Hudson, P. J. *J. Anim. Ecol.* **69**, 620–638 (2000).
3. Moran, P. A. P. *Aust. J. Zool.* **1**, 291–298 (1953).
4. Sokal, R. R. & Rohlf, F. J. *Biometry: The Principles and Practice of Statistics in Biological Research* (Freeman, New York, 1995).
5. Holyoak, M. & Lawler, S. P. *J. Anim. Ecol.* **65**, 640–652 (1996).
6. Ims, R. A. & Steen, H. *Oikos* **57**, 381–387 (1990).
7. Buonaccorsi, J. P. *et al. J. Theor. Biol.* **224**, 107–114 (2003).
8. Forchhammer, M. C. & Boertmann, D. M. *Ecography* **16**, 299–308 (1993).
9. Koenig, W. D. & Knopps, J. M. H. *Am. Nat.* **155**, 59–69 (2000).
10. Wallis, W. A. & Roberts, H. V. *Statistics: A New Approach* (Free, New York, 1956).

corrigendum

Fat-1 mice convert n-6 to n-3 fatty acids

Jing X. Kang, Jingdong Wang, Lin Wu, Zhao B. Kang
Nature **427**, 504 (2004)

J. X. K. is the inventor and co-applicant (with Massachusetts General Hospital, a non-profit organization) of a patent application relevant to this work (USN 60/275,222; WO02072028), which should therefore have been declared as a competing financial interest.

Grain boundaries as reservoirs of incompatible elements in the Earth's mantle

Takehiko Hiraga^{1*}, Ian M. Anderson² & David L. Kohlstedt¹

¹Department of Geology and Geophysics, University of Minnesota, Minneapolis, Minnesota 55455, USA

²Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA

* Present address: Graduate School of Engineering, Tohoku University, Sendai 980-8579, Japan

The concentrations and locations of elements that strongly partition into the fluid phase in rocks provide essential constraints on geochemical and geodynamical processes in Earth's interior. A fundamental question remains, however, as to where these incompatible elements reside before formation of the fluid phase. Here we show that partitioning of calcium between the grain interiors and grain boundaries of olivine in natural and synthetic olivine-rich aggregates follows a thermodynamic model for equilibrium grain-boundary segregation. The model predicts that grain boundaries can be the primary storage sites for elements with large ionic radius—that is, incompatible elements in the Earth's mantle. This observation provides a mechanism for the selective extraction of these elements and gives a framework for interpreting geochemical signatures in mantle rocks.

Segregation of incompatible elements results in a marked difference between the composition of the interiors and the boundaries of grains within the rocks that comprise the Earth's mantle¹. At chemical equilibrium, the molar concentrations of solutes in grain matrices, X_{GM} , and the grain boundaries, X_{GB} , can be related through the grain-boundary segregation energy, γ :

$$\frac{X_{GB}}{X_{GB0} - X_{GB}} = \frac{X_{GM}}{1 - X_{GM}} \exp\left(\frac{\gamma}{RT}\right) \quad (1)$$

where X_{GB0} is the saturation level of X_{GB} , R is the gas constant, and T is the temperature². This thermodynamic model has not been applied to grain boundaries in Earth materials; grain boundaries have simply been treated as possible sinks for impurities in rocks. Many studies that have inferred the presence of trace elements at grain boundaries in rocks attributed their observations to secondary effects such as contamination in intergranular regions introduced by a melt, an aqueous fluid or a gas phase^{3–10}.

To test whether this thermodynamic model describes segregation to grain boundaries in mantle rocks, we fabricated aggregates of olivine, olivine + anorthite (10 vol.%), and olivine + diopside (10 vol.%). Powders of San Carlos olivine ($Mg_{1.8}Fe_{0.2}SiO_4$) with or without synthetic anorthite or synthetic diopside glass (crystallized at 1,373 K) were isostatically hot-pressed at 1,473 K and 300 MPa for 4–7 h in a gas-medium apparatus. The aggregates were then annealed at 1,373, 1,473 or 1,523 K for ~250 h at a total pressure of one atmosphere with the oxygen fugacity fixed near the Ni/NiO buffer. Finally, the samples were quenched in water. Cracks formed owing to thermal shock, especially along some of the grain boundaries; however, many boundaries remained closed. The 1–20 μ m grains are polygonal in shape. A very small amount of melt (<1 vol.%) formed in some samples. By using different mineral assemblages and annealing temperatures, we obtained a series of aggregates with different concentrations of Ca in the olivine grains as determined by electron probe microanalysis (EPMA).

Grain-boundary compositional profiles were acquired by STEM/EDX (scanning transmission electron microscopy/energy dispersive X-ray) spectrum profiling, as detailed elsewhere^{1,11}. The incident electron probe comprised ~1.5 nA of 200-keV electrons in an incident probe of ~1.4 nm full-width at half-maximum (FWHM). For each boundary, multiple profiles were acquired in

succession and visually inspected to ensure the absence of beam damage. These profiles were then summed to obtain composite profiles with a total acquisition time of 2.5 s per pixel. Characteristic X-ray intensities were calculated by fitting reference spectra for families of elemental lines to these composite spectra.

A grain-boundary profile characteristic of the olivine + anorthite sample is shown in Fig. 1a. All profiles exhibit segregation of Ca and concomitant depletion of Mg at the olivine grain boundaries. In contrast, the prominent Al segregation shown in Fig. 1a was observed only in the olivine + anorthite specimen. Similar enrichments at olivine grain boundaries have been reported previously^{1,7,11–13}. The width of the peaks (~5 nm FWHM) is characteristic of electron-beam spreading in the specimen; smaller widths were observed for data acquired from thinner regions of the specimen. Also, the change in composition at the grain boundary does not indicate the presence of a second phase. Thin intergranular amorphous films with thicknesses of 0.5 to 10 nm observed in some polycrystalline oxides are not present at olivine grain boundaries in these specimens, as demonstrated by the lattice fringe image using high-resolution electron microscopy (HREM) in Fig. 1b (ref. 11).

Concentrations of Ca were extracted from measurements of characteristic X-ray counts above background using the standard Cliff–Lorimer¹⁴ approach. For this method, sensitivity factors for the major elements were determined using the matrix as a standard, with the composition as determined by EPMA. Comparable sensitivity factors for Ca were taken as the mean of the factors for Si and Fe, effectively a linear interpolation of the factors for these major elements versus atomic number. The accuracy of this interpolation, which yielded $k_{CaSi} = 1.17 \pm 0.08$ for all measurements, was deemed adequate given the statistical limitations of the measurement of the low Ca X-ray intensities. To compare the degree of segregation among specimens, the total excess Ca integrated through the width of the profile was determined and converted to effective monolayers at the boundary. The procedure for this calculation is well described in ref. 1. Because the profiles indicate that Ca substitutes for Mg at the boundary, the grain-boundary coverage of Ca is expressed as molar occupancy of octahedral M (M1 and M2) regular lattice sites, X_{GB}^{Ca} . This measure provides an upper bound for the grain-boundary concentration, because it

represents the concentration if the solute were confined entirely to the grain-boundary plane. This concentration may, in fact, represent the real grain-boundary concentration, as observed for segregation of Ca to within a single atomic plane at grain boundaries in Ca-doped MgO (ref. 15).

In Fig. 2, Ca monolayer coverage is plotted versus grain-matrix Ca concentration (molar occupancy of octahedral M sites), X_{GM}^{Ca} , for the samples described above combined with the results from samples of two natural peridotites, a fine-grained lherzolite and an olivine phenocryst-bearing basaltic rock, as well as a synthetic olivine + diopside aggregate¹. Individual data points represent the coverage from a single grain boundary in the samples. The observed variation in Ca coverage among grain boundaries in a given sample can be explained by the variation of grain-boundary structure for different misorientations between neighbouring grains and grain-boundary orientation. However, the plots demonstrate that (1) the concentration of Ca at the grain boundaries increases with increasing concentration in the olivine grains, (2) the

partition coefficient, $D_{GM/GB}^{Ca} = X_{GM}^{Ca} / X_{GB}^{Ca}$, decreases with decreasing temperature, and (3) $D_{GM/GB}^{Ca}$ increases with increasing Ca concentration in the matrix. All of these features are predicted by equation (1).

Strain energy minimization and space charge compensation are two important thermodynamic driving forces for segregation to grain boundaries¹⁶. A survey of all elements measured in the grain-boundary profiles indicates an enrichment of Ca, K, Cr, Ti, Al and occasionally Fe, depletion of Mg and occasionally Si, and no detectable enrichment or depletion of Mn and Ni at the grain boundaries^{1,11} (T.H., unpublished work). These observations suggest that differences in ionic radius and valence state from that of host cations (that is, Mg and Si) control the enrichment of elements at grain boundaries. Divalent cations such as Ca^{2+} , which are enriched at grain boundaries, have much larger ionic radii (0.100 nm) than those of the host isovalent cations ($Mg^{2+} = 0.072$ nm and $Fe^{2+} = 0.078$ nm). In contrast, Mn^{2+} and Ni^{2+} , which do not segregate, have ionic radii ($Mn^{2+} = 0.083$ nm and $Ni^{2+} = 0.069$ nm) close to those of the host cations. For these isovalent cations, we conclude that segregation is controlled by differences in ionic radius. Similar partitioning behaviour takes place between coexisting minerals (for example, olivine, pyroxenes and plagioclase) and between minerals and melt^{17,18}.

Aliovalent elements such as K, Cr, Al and Ti can segregate to grain boundaries to compensate space charge. Al commonly segregates to olivine grain boundaries and occasionally exceeds the amount of Ca, as shown in Fig. 1a. Because Si and Fe concentrations are constant in the profile, Al must substitute for Mg at the grain boundaries. The solubility of Al in olivine grains is below the detection limit of EPMA, indicating that it partitions more strongly to grain boundaries than does Ca, even though the ionic radius of Al^{3+} (0.054 nm) is closer than that of Ca to the value of either the radii of host cations or the theoretical strain-free radius discussed later. Likewise the ionic radius of Cr^{3+} (0.062 nm) is approximately equal to the theoretical strain-free radius for M2 sites in olivine, yet segregation to olivine grain boundaries was detected¹. These observations indicate that aliovalency is also a strong driving factor for partitioning to grain boundaries. However, the combination of strain and valence factors complicates the development of a quantitative model for predicting the partitioning of aliovalent ions.

The partition coefficients for isovalent elements are well explained by the strain energy, U , caused by the misfit of trace

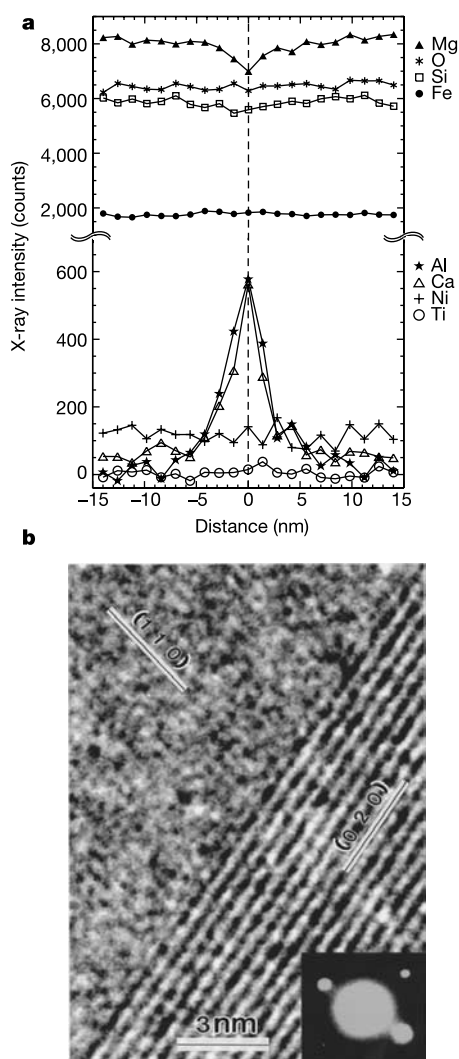


Figure 1 Chemistry and structure of olivine-olivine grain boundaries in a sample of olivine + anorthite annealed at 1,473 K. **a**, X-ray intensity profile from STEM/EDX analysis for elements in the vicinity of the grain boundaries. The profile is the sum of profiles measured from five grain boundaries. The negative compositional values of trace elements with low concentrations in the olivine grains are indicative of the statistical scatter in the spectral-fitting procedure. **b**, HREM image and diffraction pattern of the olivine grain-boundary area.

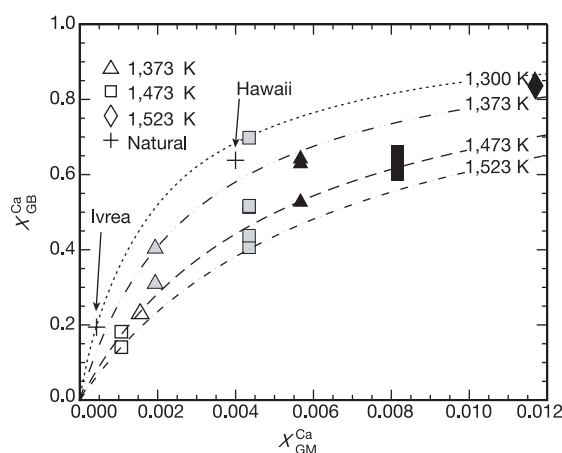


Figure 2 Calcium concentrations in olivine grain boundaries versus concentration in olivine grain matrices. Open symbols, olivine only; grey symbols, olivine + anorthite; black symbols, olivine + diopside. Theoretical curves are shown for four temperatures from equations (1) and (2).

elements in the lattice¹⁹

$$U = 4\pi E_{\alpha} N_A \left[\frac{r_0}{2} (r_i - r_0)^2 + \frac{1}{3} (r_i - r_0)^3 \right] \quad (2)$$

where E_{α} is the effective Young's modulus of the lattice site α , N_A is Avogadro's number, r_i is the ionic radius of the substitution cation, and r_0 is the radius of the unoccupied site of interest, in other words, the theoretical strain-free radius. Studies of the olivine-melt system indicate that E_{α} is similar to that of the bulk, so we used previously measured values for the Young's modulus of olivine (Fo₉₀) normalized to the temperature conditions of our experiments²⁰. We calculated r_0 from bond lengths of M2 octahedron in forsterite²¹, $r_0 = r_{M-O} - 0.5 \times r_{O-O}$, where r_{M-O} and r_{O-O} are the average spacings between M and oxygen sites and oxygen and oxygen sites, respectively. Values of r_0 for different phases and as a function of temperature are summarized in Table 1. The value of r_0 is similar to that obtained by fitting the observed partitioning data in Fig. 1 to equations (1) and (2), the procedure used to determine r_0 in a study of partitioning between olivine and melt²². If the misfit strain energy is relieved for solutes that segregate to grain boundaries, the grain-boundary absorption energy for divalent elements is likely to be dominated by the misfit strain energy. For various oxides, grain-boundary segregation is well explained by the misfit energy²³.

We calculated partitioning curves for Ca ($r_i = 0.100$ nm) using equation (1) for different temperatures, assuming that the segregation energy, γ , is equal to misfit strain energy, U , and that the segregated Ca is confined to the M sites at the grain-boundary plane ($X_{GB0} = 1$). These calculated curves are also shown in Fig. 2. Although some deviations exist between the theoretical curves and experimental partitioning data, the good overall agreement suggests that the observed Ca partitioning is consistent with the thermodynamic equilibrium model in equation (1), with the segregation confined to the grain-boundary plane and the partitioning controlled by misfit strain energy. The partitioning of Ca to boundaries in two natural samples is stronger than that observed for synthetic samples, possibly because the natural samples equilibrated at lower temperature, $\sim 1,300$ K, which is within the range estimated for the equilibration temperature for the Ivrea sample²⁴. The consistency of the Ca partitioning in both natural and synthetic samples with our model predictions suggests that partitioning according to this segregation model accurately describes the partitioning of Ca, and probably other incompatible divalent cations, in the Earth's mantle.

In previous geochemical investigations, grain boundaries have not been treated as sites for the accommodation of various elements, according to the laws of thermodynamics, as we have done here. Accordingly, the distribution of elements in rocks has been interpreted on the basis of their solubility in the various distinct mineral and fluid phases. Although the number of sites at the grain-boundary planes is small in comparison with that in the grain matrix ($\sim 10^{-6}$ for a grain size of 1 mm), the small value of $D_{GM/GB}$ suggests that grain boundaries may store significant amounts of incompatible elements. This finding may answer a fundamental question in mantle geochemistry: where are incompatible elements stored in rock, especially fluid-free rock?

Table 1 Parameters used in this study

	T (K)	$N_M (\times 10^{22} \text{ cm}^{-3})$	r_0 (nm)	E_{α} (GPa)
Olivine	1,373	2.69	0.064	159
Olivine	1,473	2.69	0.065	155
Olivine	1,523	2.69	0.065	154
Clinopyroxene	1,473	1.83	0.105*	182*
Orthopyroxene	1,473	2.68	0.078†	147†

* Values for element partitioning between diopside and melt from ref. 18.

† Values for element partitioning between orthopyroxene and melt from ref. 22.

To evaluate the storage capacity of grain boundaries for incompatible elements, we calculated the ratio of solute concentration (weight fraction) in a system composed of grain matrices + grain boundaries, C_{GM+GB} , relative to that of a system composed of only grain matrices, C_{GM} , for the elements with large ionic radii. This ratio can be compared to that from a bulk-rock analysis normalized by the rock composition calculated from the compositions of the constituent mineral phases and mineral modes as described previously^{6,8,10}. The ratio C_{GM+GB}/C_{GM} can be expressed as

$$\frac{C_{GM+GB}}{C_{GM}} \approx \frac{X_{GB} V_{GB} + X_{GM} V_{GM}}{X_{GM} (V_{GB} + V_{GM})} \quad (3)$$

where V_{GB} and V_{GM} are the volumes of grain boundaries and grain matrices, respectively. In this calculation, we ignored the change of density in V_{GB} and V_{GM} due to substitution of different elements, because this factor is relatively insignificant in a plot containing the logarithm of C_{GM+GB}/C_{GM} . For tetrakaidecahedral grains, $V_{GB} = 1.73 \times t \times d^2$ and $V_{GM} = 0.52 \times d^3$, where t is thickness of the boundary and d is the grain size²⁵. For the monolayer approximation, t is given by $(N_M)^{-1/3} \times 10^7$ nm, where N_M is defined as the volume concentration of M (M1 and M2) sites (Table 1). For dilute solutions, X_{GB}/X_{GM} can be approximated as $\exp(\gamma/RT)$, from equation (1), again only considering the effects of misfit strain energy (that is, $\gamma = U$) from the ionic radii of solutes at constant temperature, as long as the radius of the unoccupied site is known (see equation (2)). This prediction has been confirmed for Al₂O₃ doped with solutes having various ionic radii²⁶.

Using equation (3) with the above approximation for X_{GB}/X_{GM} , theoretical curves for C_{GM+GB}/C_{GM} versus ionic radius can be plotted for olivine aggregates of different grain sizes, as shown in Fig. 3. In reality, upper mantle rocks are composed of orthopyroxene (5–30%), clinopyroxene (5–20%) and a small amount of spinel ($\leq 3\%$), in addition to olivine (60–80%). In general, trace elements partition most strongly into clinopyroxene and more strongly into orthopyroxene than into olivine. This partitioning behaviour can be

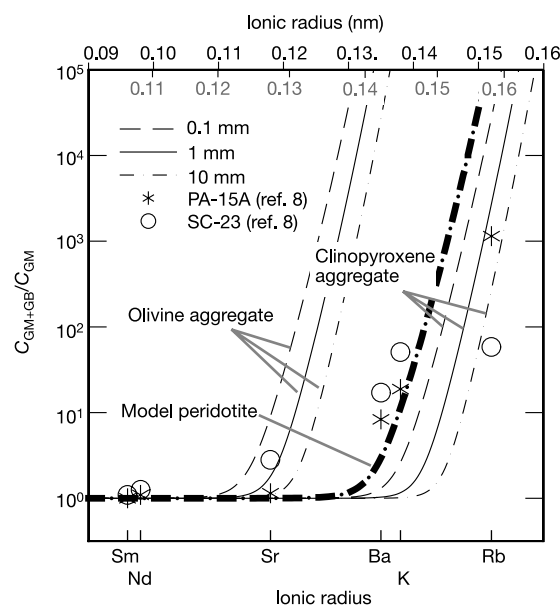


Figure 3 Ratio of solute concentration in a system composed of grain matrices + grain boundaries, C_{GM+GB} , to that of a system composed of grain matrices alone, C_{GM} , versus ionic radius at a temperature of 1,473 K. Ionic radii for both sixfold coordination (olivine) and eightfold coordination (clinopyroxene) are shown as upper black values and lower grey values, respectively. Plots are C_{GM+GB}/C_{GM} observed in real rock samples (PA-15A and SC-23) reported in ref. 8. The predicted curve for model peridotite of SC-23 is calculated from equation (3) with reported values of mineral modes, grain size and temperature from refs 8 and 27.

explained in terms of the relative values of r_0 for these minerals. Thus, X_{GB}/X_{GM} will be largest for olivine and smallest for clinopyroxene for any specific value of ionic radius larger than ~ 0.1 nm. Consequently, we extended our model of the grain-boundary partitioning to include clinopyroxene in order to estimate the limits of grain-boundary segregation in mantle rocks. To obtain the ratio C_{GM+GB}/C_{GM} for a clinopyroxene aggregate, we used equation (2) with the approximation $X_{GB}/X_{GM} \approx \exp(U/RT)$ and the parameters for diopside ($\text{CaMgSi}_2\text{O}_6$) given in Table 1. We calculated the ratio C_{GM+GB}/C_{GM} for the aggregates with grain sizes of 0.1, 1 and 10 nm, a range that encompasses the grain size of most mantle rocks. We assumed an equilibration temperature of 1,473 K to compare with previously reported geochemical data. The M2 site in clinopyroxene is eightfold coordinated rather than sixfold as in olivine, so for each element in clinopyroxene we used an ionic radius different from that used for the same element in olivine. We also assumed that the misfit strain energy is entirely relieved at phase boundaries (for example, olivine/clinopyroxene); therefore, curves of C_{GM+GB}/C_{GM} versus ionic radius for olivine–orthopyroxene–clinopyroxene aggregates lie between those for pure olivine and pure clinopyroxene aggregates.

Because the fraction of sites at the grain boundary increases as grain size decreases, grain-boundary segregation has a more pronounced effect on bulk chemistry at smaller ionic radius for fine-grained rocks. Grain boundaries start to influence the bulk chemistry for incompatible elements with ionic radii greater than approximately 0.11 nm depending on grain size and constituent minerals (Fig. 3). As temperature decreases, the curves shift towards smaller ionic radii, resulting in a decrease of the threshold from 0.11 to ~ 0.10 nm for a pure olivine aggregate at 973 K. Similarly, as temperature increases, the curves shift towards larger ionic radii resulting in an increase of the threshold to ~ 0.12 nm at 1,873 K. For aliovalent ions such as rare-earth elements, for which space charge compensation is a significant driving force for grain-boundary segregation, the curves of C_{GM+GB}/C_{GM} versus ionic radius in Fig. 3 will shift towards smaller ionic radii.

Geochemical studies have often presumed that significant amounts of incompatible elements reside in intergranular regions, because the bulk-rock chemistry differs from that of pulverized rocks washed in acid to remove the surfaces of the grains and/or that of ‘reconstituted’ rocks obtained from the composition and modal abundances of the mineral phases^{3,4,6,8–10}. However, all of these studies concluded that their observations result from secondary effects associated with contamination, such as may be introduced by ground water. Menzies and Murthy³ reported that acid washing dissolves intergranular material, thus lowering the concentrations of K, Rb and Ba in their lherzolite sample, while not affecting the concentration of Sr. Zindler and Jagoutz⁸ found strong positive correlations between ionic radii and trace element concentrations in a bulk rock normalized by that in an acid-cleaned rock or a ‘reconstituted’ rock for ionic sizes of Sr or larger in their samples, as shown in Fig. 3. We calculated a curve for their rock sample SC-23 (ref. 8) using their reported values of grain size, mineral modes and temperature²⁷. The ionic radius of Sr falls between 0.11 and 0.12 nm, approximately dividing the regime of whole-rock chemistry controlled by olivine grain matrices from that controlled by olivine grain boundaries. The ratios C_{GM+GB}/C_{GM} from the data of ref. 8 start to increase with increasing ionic radius at an ionic size slightly smaller than predicted by our calculation; also, the rate of increase differs somewhat from that expected from our calculation. However, overall, the trends are comparable to those anticipated from our model. The deviation can be due to inhomogeneity of the rock texture associated with grain size and shape and to uncertainties in the effect of temperature and pressure on the parameters that we used (Table 1).

In mantle peridotites, a similar correlation between ionic radii and the concentrations of rare-earth elements with the same

valences occurs in a bulk rock normalized by a ‘reconstituted’ rock^{4,6,10}. This behaviour was also attributed to secondary contamination; however, equilibrium rare-earth-element segregation driven by lattice strain energy and space charge must be a primary cause of the observed rare-earth-element distributions.

The majority of Earth’s mantle is fluid-free. Mantle, especially primitive mantle, supplies incompatible elements to fluids that bring these elements to the surface. However, the major constituents of mantle rocks do not contain enough potassium and other incompatible elements to explain the abundance or the distribution of these elements in basalts²⁸. Many studies have proposed that the primary storage sites for these elements are accessory minerals²⁹, but this point is much debated. From our experimental observations combined with calculations based on equations (1) and (2), we conclude that grain boundaries can be the primary storage sites for incompatible elements. Preliminary experiments demonstrate that K partitions strongly to olivine grain boundaries (T.H., unpublished observations).

At the initial stage of melting, melts are produced and migrate primarily along grain boundaries and grain edges³⁰. Geochemical models for this stage in the mantle, for example beneath mid-ocean ridges, are based on the element partition coefficients between melt and crystal, in addition to kinetic factors such as the diffusivities of the elements in the crystalline phases³¹. However, the selective melting of intergranular regions that store high concentrations of incompatible elements such as U and Th, which have been key elements used to estimate the production of mid-ocean-ridge basalt from primitive mantle³², provides an alternative mechanism for the selective extraction of these elements. Further investigation of $D_{GM/GB}$ for these elements is necessary.

If volatile fluids and low-viscosity melts infiltrate through formerly fluid-free rocks driven by the reduction in interfacial energy, for example, selective dissolution would preferentially occur at grain edges³³. The concentrations of incompatible elements in these types of fluids will be out of equilibrium with those in the crystalline grains due to fast fluid migration and slow ionic diffusion. As a result, the fluids will be enriched in those elements that strongly partition to grain boundaries. This process might also explain the production of metasomatic fluids (for example, carbonatite-like liquids), which are often observed to be highly enriched in incompatible elements in mantle wedges^{34,35}.

Here we have shown that grain boundaries store incompatible elements with concentrations that are thermodynamically controlled. The common behaviour of synthetic and natural samples demonstrates that laboratory experiments can accurately model partitioning behaviour in the Earth’s mantle. We have proposed some possible consequences of our results for our understanding of fundamental geochemical processes in the Earth’s mantle. However, the consequences of grain boundaries acting as primary storage sites for incompatible elements are likely to be much broader. Chemical segregation also affects grain-boundary processes such as grain-boundary diffusion, sliding, fracture and migration, all of which influence the physical properties of the Earth’s mantle. □

Received 24 April; accepted 27 November 2003; doi:10.1038/nature02259.

- Hiraga, T., Anderson, I. M. & Kohlstedt, D. L. Chemistry of grain boundaries in mantle rocks. *Am. Mineral.* **88**, 1015–1019 (2003).
- McLean, D. *Grain Boundaries in Metals* 346 (Clarendon, Oxford, 1957).
- Menzies, M. M. & Murthy, V. R. Strontium isotope geochemistry of alpine tectonite lherzolites: Data compatible with a mantle origin. *Earth Planet. Sci. Lett.* **38**, 346–354 (1978).
- Ottone, G. Rare earth abundances and distribution in some spinel peridotite xenoliths from Assab (Ethiopia). *Geochim. Cosmochim. Acta* **44**, 1885–1901 (1980).
- Fraser, D. G., Watt, F., Grime, G. W. & Takacs, J. Direct determination of strontium enrichment on grain boundaries in a garnet lherzolite xenolith by proton microprobe analysis. *Nature* **312**, 352–354 (1984).
- Stosch, H.-G., Lugmair, G. W. & Kovalenko, V. I. Spinel peridotite xenoliths from the Tariat Depression, Mongolia. II: Geochemistry and Nd and Sr isotopic composition and their implications for the evolution of the subcontinental lithosphere. *Geochim. Cosmochim. Acta* **50**, 2601–2614 (1986).
- Suzuki, K. Grain-boundary enrichment of incompatible elements in some mantle peridotites. *Chem. Geol.* **63**, 319–334 (1987).

8. Zindler, A. & Jagoutz, E. Mantle cryptology. *Geochim. Cosmochim. Acta* **52**, 319–333 (1988).
9. Ionov, D. A., Kramm, U. & Stosch, H.-G. Evolution of the upper mantle beneath the southern Baikal rift zone: an Sr–Nd isotope study of xenoliths from the Bartoy volcanoes. *Contrib. Mineral. Petrol.* **111**, 235–247 (1992).
10. Xu, X., O'Reilly, S. Y., Griffin, W. L. & Zhou, X. Enrichment of upper mantle peridotite: petrological, trace element and isotopic evidence in xenoliths from SE China. *Chem. Geol.* **198**, 163–188 (2003).
11. Hiraga, T., Anderson, I. M., Zimmerman, M. E., Mei, S. & Kohlstedt, D. L. Structure and chemistry of grain boundaries in deformed, olivine + basalt and partially molten lherzolite aggregates: Evidence of melt-free grain boundaries. *Contrib. Mineral. Petrol.* **144**, 163–175 (2002).
12. Waff, H. S. & Holdren, G. R. The nature of grain boundaries in dunite and lherzolite xenoliths: Implications for magma transport in refractory upper mantle material. *J. Geophys. Res.* **86**, 3677–3683 (1981).
13. Tan, B. H., Jackson, I. & Fitz Gerald, J. D. High-temperature viscoelasticity of fine-grained polycrystalline olivine. *Phys. Chem. Miner.* **28**, 641–664 (2001).
14. Cliff, G. & Lorimer, G. W. The quantitative analysis of thin specimens. *J. Microsc.* **103**, 203–207 (1975).
15. Yan, Y., Chisholm, M. F., Duscher, G. & Pennycook, S. J. Atomic structure of a Ca-doped [001] tilt grain boundary in MgO. *Electron Microsc.* **47**, 115–120 (1998).
16. Kingery, W. D. Plausible concepts necessary and sufficient for interpretation of ceramic grain-boundary phenomena. II. Solute segregation, grain-boundary diffusion, and general discussion. *J. Am. Ceram. Soc.* **57**, 74–83 (1974).
17. Beattie, P. Systematics and energetics of trace-element partitioning between olivine and silicate melts: Implications for the nature of mineral/melt partitioning. *Chem. Geol.* **117**, 57–71 (1994).
18. Blundy, J. & Wood, B. Prediction of crystal-melt partition coefficients from elastic moduli. *Nature* **372**, 452–454 (1994).
19. Brice, J. C. Some thermodynamic aspects of the growth of strained crystals. *J. Cryst. Growth* **28**, 249–253 (1975).
20. Isaak, D. G. High temperature elasticity of iron-bearing olivines. *J. Geophys. Res.* **97**, 1871–1885 (1992).
21. Takeuchi, Y., Yamanaka, T., Haga, N. & Hirana, M. in *Material Science of the Earth's Interior* (ed. Sunagawa, I.) 191–231 (Terra Scientific, Tokyo, 1984).
22. Purton, J. A., Allan, N. L., Blundy, J. D. & Wasserman, E. A. Isovalent trace element partitioning between minerals and melts: A computer simulation study. *Geochim. Cosmochim. Acta* **60**, 4977–4987 (1996).
23. Johnson, W. C. Grain boundary segregation in ceramics. *Metall. Trans. A* **8A**, 1413–1422 (1977).
24. Shervais, J. Thermal emplacement model for the alpine lherzolite massif at Balmuccia (Italy). *J. Petrol.* **20**, 795–820 (1979).
25. German, R. M. Formation of necklace microstructures during liquid phase sintering: model calculation. *Int. J. Powder Metall.* **22**, 31–38 (1986).
26. Li, C.-W. & Kingery, W. D. in *Advances in Ceramics* (ed. Kingery, W. D.) 368–378 (American Ceramic Society, Columbus, 1985).
27. Frey, F. A. & Prinz, M. Ultramafic inclusions from San Carlos, Arizona: Petrologic and geochemical data bearing on their petrogenesis. *Earth Planet. Sci. Lett.* **38**, 129–176 (1978).
28. Oxburgh, E. R. Petrological evidence for the presence of amphiboles in the upper mantle and its petrogenetic and geophysical implications. *Geol. Mag.* **101**, 1–11 (1964).
29. Wang, W. & Takahashi, E. J. Subsolidus and melting experiments of K-doped peridotite KLB-1 to 27 GPa: Its geophysical and geochemical implications. *Geophys. Res.* **105**, 2855–2868 (2000).
30. Watson, E. B., Brennan, J. M. & Baker, D. R. in *Continental Mantle* (ed. Menzies, M.) 111–125 (Oxford Univ. Press, Oxford, 1991).
31. Van Orman, J. A., Grove, T. L. & Shimizu, N. Uranium and thorium diffusion in diopside. *Earth Planet. Sci. Lett.* **160**, 505–519 (1998).
32. Hauri, E., Wagner, T. P. & Grove, T. L. Experimental and natural partitioning of Th, U, Pb and other trace elements between garnet, clinopyroxene and basaltic melts. *Chem. Geol.* **117**, 149–166 (1994).
33. Nakamura, M. & Watson, E. B. Experimental study of aqueous fluid infiltration into quartzite: implications for the kinetics of fluid redistribution and grain growth driven by interfacial energy reduction. *Geofluids* **1**, 73–89 (2001).
34. Maury, R. C., Defant, M. J. & Joron, J. L. Metasomatism of the sub-arc mantle inferred from trace elements in Philippine xenoliths. *Nature* **360**, 661–663 (1992).
35. Zanetti, A., Mazzucchelli, M., Rivalenti, G. & Vannucchi, R. The Finero phlogopite-peridotite massif: an example of subduction-related metasomatism. *Contrib. Mineral. Petrol.* **134**, 107–122 (1999).

Acknowledgements We thank M. Hirschmann and T. Morishita for discussions and M. Zimmerman for experimental assistance. T.H. acknowledges receipt of a JSPS research fellowship. This collaboration was performed under the ORNL SHaRE User Program and with support from the NSF.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T.H. (hirag001@umn.edu).

Whisker movements evoked by stimulation of single pyramidal cells in rat motor cortex

Michael Brecht¹, Miriam Schneider¹, Bert Sakmann¹ & Troy W. Margrie^{1,2}

¹Department of Cell Physiology, Max Planck Institute for Medical Research, Jahnstrasse 29, D-69120 Heidelberg, Germany

²Department of Physiology, The Wolfson Institute for Biomedical Research, University College London, Gower Street, London WC1E 6BT, UK

Neuronal activity in the motor cortex is understood to be correlated with movements, but the impact of action potentials (APs) in single cortical neurons on the generation of movement has not been fully determined. Here we show that trains of APs in single pyramidal cells of rat motor cortex can evoke long sequences of small whisker movements. For layer-5 pyramids, we find that evoked rhythmic movements have a constant phase relative to the AP train, indicating that single layer-5 pyramids can reset the rhythm of whisker movements. Action potentials evoked in layer-6 pyramids can generate bursts of rhythmic whisking, with a variable phase of movements relative to the AP train. An increasing number of APs decreases the latency to onset of movement, whereas AP frequency determines movement direction and amplitude. We find that the efficacy of cortical APs in evoking whisker movements is not dependent on background cortical activity and is greatly enhanced in waking rats. We conclude that in vibrissae motor cortex sparse AP activity can evoke movements.

In mammals, a key brain structure in motor control is the primary motor cortex (M1), a large frontal cortical area^{1–3}. The role of M1 neurons in generating movements has been explored by extracellular stimulation with small currents (intracortical microstimulation)⁴ and by extracellular recordings of single neurons^{5,6}.

Extracellular stimulation experiments in M1 have shown that there are very low current thresholds (as low as 1–2 μ A) for movement initiation³ that are thought to stimulate few cortical neurons^{7,8}. Extracellular recordings from single cells in M1 of awake, behaving monkeys have also shown that there is a close match between AP activity and movement⁶. The quantitative analysis of such recordings by an AP population vector^{9,10} and by real-time analysis of AP activity in M1—motivated by the desire to construct cortically controlled neuroprosthetic devices^{11–14}—has led to the conclusion that the activity of only a few hundred to a few tens of neurons is sufficient to specify accurately complex limb movements.

The relationship between neuronal and muscular activity has been also analysed by computing averages of muscle activity (assayed by electromyographic recordings) triggered relative to neuronal APs¹⁵. This technique has revealed significant correlations between cortical single-cell activity and muscle activity^{15–17}. Thus, extracellular stimulation and recording studies indicate that there is a close correlation between activity in the motor cortex and muscular activity, consistent with the known synaptic connectivity between cortex and motoneurons¹⁶. Intracellular stimulation of cortical cells evokes changes in electromyographic recordings¹⁸ and might allow a more direct investigation of the relationship between cortical APs and movement.

In rats, whisker movements can escape from sleep or anaesthesia-induced paralysis^{19,20}. Such movements are rhythmic, are restricted to the plane of the whisker row, and can be quantified as a simple back-and-forth movement^{21–24}. The movements of individual whiskers can be tracked readily, making them a convenient system to quantitatively study the generation of motor output²⁵. Here we have used whole-cell recordings to examine the motor effects of intracellularly evoked APs in identified layer-5 (L5) and layer-6 (L6) pyramidal cells in the vibrissae motor cortex²⁶ of anaesthetized rats.

Movements evoked by intracellular stimulation

Figure 1 shows a pyramidal L6 cell (Fig. 1a) in which intracellular current injection evoked 10 APs at 50 Hz (Fig. 1c). This stimulation caused backward movements of whisker E1 and adjacent whiskers (Fig. 1b and Supplementary Movie 1), which were the same whiskers that were deflected backwards by extracellular stimulation in the same region (Fig. 1d). Movements consisted of bouts of backwards-and-forwards sweeps, which we refer to as a ‘whisking burst’, beginning after a variable latency of 0.05–2.0 s. The peak of the evoked movement was delayed (0.3 s to 50% of the amplitude) and the movement lasted for several seconds (Fig. 1e). We quantified the peak-to-peak amplitude of movements in the second immediately after current injection and in the second just before current injection. Action potential initiation evoked significant movements ($2.12 \pm 0.17^\circ$ poststimulation amplitude versus $1.62 \pm 0.14^\circ$ prestimulation amplitude; paired *t*-test, $P < 0.05$) but injection of hyperpolarizing currents did not (paired *t*-test, $P > 0.3$; Fig. 1e), resulting in significantly larger ratios of post- to prestimulation amplitudes for AP initiation than for hyperpolarizing current injection (paired *t*-test, $P < 0.05$).

To assess the significance of movement effects in individual experiments, we measured the movement responses trial by trial across 10–30 stimulations. Peak-to-peak amplitudes of whisker movements in the second just after initiation of APs were compared with peak-to-peak amplitudes of spontaneous movements in the second immediately before stimulation. A significant poststimulation increase in whisker movements was detected in 9 out of 44 regularly spiking or intrinsically bursting neurons (two-tailed, paired *t*-tests, $P < 0.05$). In addition, we analysed the degree to which averaged movement records in the first second after AP initiation ($n = 44$ cells) differed in peak-to-peak amplitude from movements occurring in the second preceding stimulation. This measurement showed that, across cells, poststimulation changes in average whisker position were significantly larger than prestimulation changes (ratio of post- to prestimulation amplitudes, 1.30 ± 0.09 ; paired *t*-test, $P < 0.01$; Fig. 2a).

In control experiments, 5-nA current steps applied through patch pipettes to the extracellular space of M1 (threefold larger than those

applied in intracellular stimulation (1.68 ± 0.8 -nA steps)) did not evoke significant whisker movements (ratio of post- to prestimulation amplitudes, 1.03 ± 0.13 ($n = 10$); Fig. 2b). Whereas AP-evoked movements were clearly stimulus locked (Fig. 2c), even the largest poststimulation deflections observed with extracellular stimulation were not (Fig. 2d). Similarly, in 13 cells hyperpolarizing currents did not evoke whisker movements (ratio of post- to prestimulation amplitudes, 1.14 ± 0.14), whereas APs evoked by depolarizing currents did (ratio of post- to prestimulation amplitudes, 1.49 ± 0.17 ; Fig. 2a, c). In five neurons of L5 and L6 of the primary visual cortex, intracellularly evoked APs did not evoke whisker movements (ratio of post- to prestimulation amplitudes, 1.02 ± 0.24 ; Fig. 2b). Thus, in none of the 31 individual control experiments with extracellular, hyperpolarizing, subthreshold current injection or AP stimulation of visual cortex neurons did we observe significant effects on movement. Prestimulation and poststimulation amplitudes did not differ between control experiments (paired t -test, $P > 0.22$), and the ratio of poststimulation to prestimulation amplitudes was significantly larger in AP initiation experiments (unpaired t -test, $P < 0.04$). Thus, the control experiments indicate that AP trains in single, deep-layer pyramidal cells in M1 cause whisker movement.

Effects of intra- and extracellular stimulation

Whisker movements evoked by intracellular stimulation were small in amplitude: most were less than 1° , and the largest were 2 – 3° after averaging. In weakly effective cells, the effect of intracellular stimulation was restricted to modulation of the ongoing whisker movements, whereas in the most effective cells stimulus-locked movements were evoked in most of the trials. Evoked movements had long latencies (minimum ~ 50 ms, but often longer) and were complex in that they always involved several whiskers and usually consisted of second-long sequences of movements (Supplementary Movies 1 and 2). The rhythmicity of evoked movements was always the same as that of ongoing movements and varied with the depth of anaesthesia.

Extracellular stimulation, at current levels just above threshold ($32 \pm 22 \mu\text{A}$), evoked short-latency (<100 ms), small-amplitude ($<5^\circ$) whisker movements. Movements evoked by either intracellular stimulation of pyramidal neurons or extracellular stimulation shared the same topography of deflected whiskers and initial direction of deflection (in 85% of cases).

Movements evoked by extracellular stimulation had a shorter latency, were more stringently locked to stimulation onset, and had a shorter duration (usually <0.5 s; Fig. 1d, e) than did movements

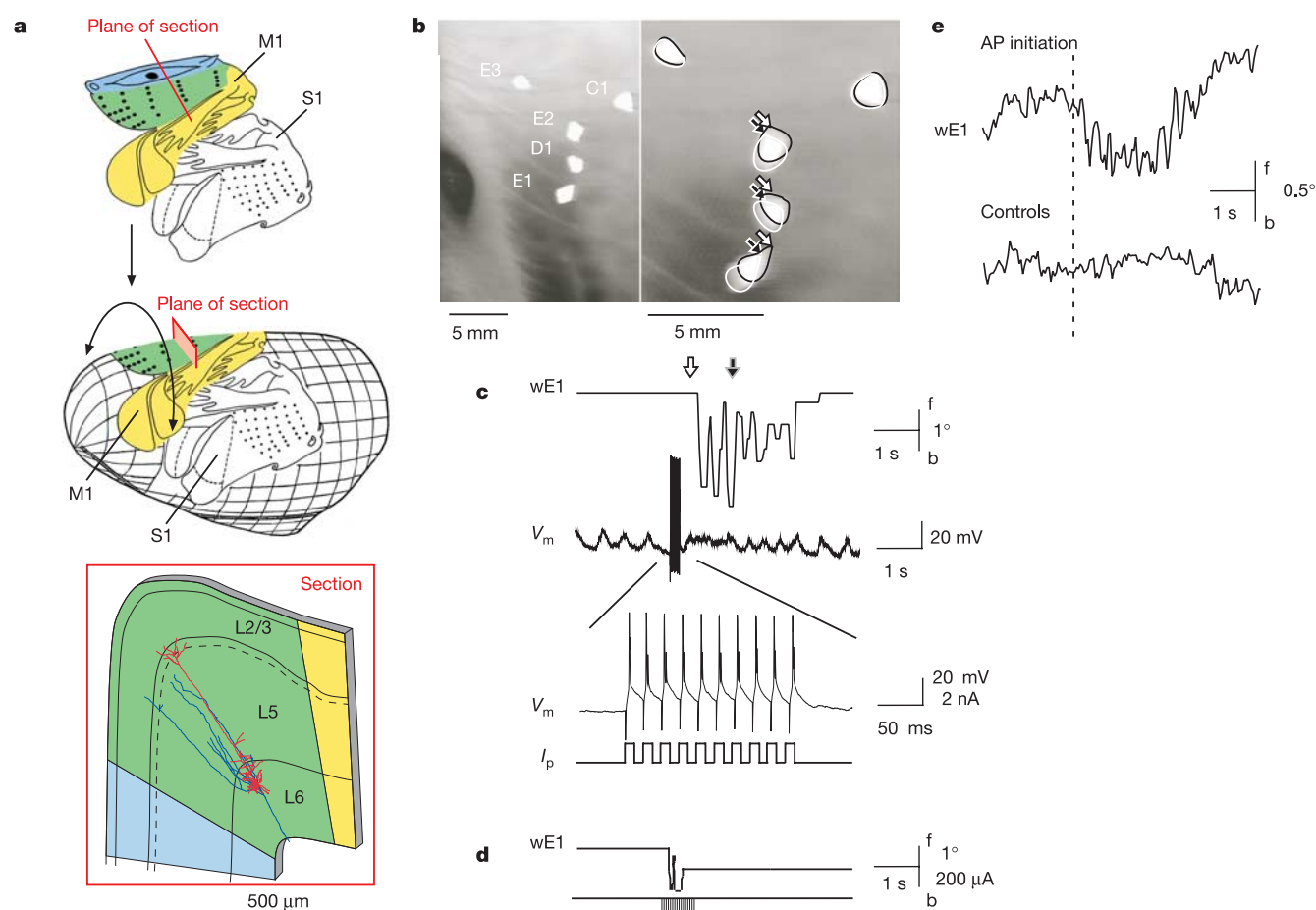


Figure 1 Whisker movements evoked by intracellular stimulation of an L6 pyramidal neuron. **a**, Top, flattened maps of the primary motor (M1) and somatosensory (S1) cortices. Middle, dorsolateral view of a rat cortical hemisphere, indicating S1 and M1. Bottom, coronal section through M1, indicating the dendrites (red) and axon (blue) of the stimulated L6 pyramidal. Blue, area Cg1 (eye movement motor cortex); green, area AGm (vibrissae motor cortex); yellow, area AGl (somatic motor cortex). **b**, Left, top view of rat head and foil-labelled (bright spots) whiskers. Right, superimposed images of prestimulation (black outlines) and poststimulation (white outlines) whisker positions

(arrows). **c**, Top trace, position of whisker E1 (wE1) in an intracellular stimulation trial (10 APs at 50 Hz). Arrows indicate the time points at which the images shown in **b** were taken. Bottom traces, membrane potential recordings and current injection steps. f, forward movement; b, backward movement. **d**, Position of whisker E1 in an extracellular stimulation trial (top) and the stimulation pattern (bottom). **e**, Movement averages of 30 single-cell stimulation trials (top) and 30 control trials (bottom) with hyperpolarizing current steps. Broken line indicates the onset of current injection.

evoked by intracellular stimulation. Unexpectedly, although single or dual whisker movements were regularly evoked by extracellular stimulation (at ~25% of sites, 14 out of 55 experiments), we never observed single or dual whisker movements after intracellular stimulation. Thus, at threshold levels, extracellular stimulation can have smaller movement fields than can intracellular stimulation.

Patterning of whisker movements is layer-specific

To understand how evoked whisker movements are related to the laminar position of motor cortex cells, we analysed movement patterns evoked by L5 ($n = 12$) and L6 ($n = 11$) pyramidal cells, in which we had observed poststimulation movements that were larger than prestimulation movements (ratio of post- to prestimulation amplitudes, >1.2). We found that stimulation of an L5 pyramidal cell (Fig. 3a) evoked movements in which the timing of an individual forwards-and-backwards whisking movement relative to the stimulation AP train was similar from one trial to the next (Fig. 3b and Supplementary Movie 2). We refer to this movement as whisker movements with a 'constant phase relative to stimulation'. As a consequence, after averaging the movement records, modulation at the whisking frequency became apparent (Fig. 3c). The power spectra of both the individual trial and the averaged movement are characterized by a peak at 3.5 Hz, which corresponds to the

whisking frequency (data not shown). In the population average, we also observed modulation of the evoked movement at the whisking frequency ($n = 12$; Supplementary Fig. 1a).

Stimulation of an L6 cell (Fig. 3d) evoked rhythmic whisker movements in which the timing of an individual forwards-and-backwards movement varied from one trial to the next (Fig. 3e and Supplementary Movie 1). Thus, the L6-evoked whisker movements had a 'variable phase relative to stimulation', and the periodic modulation at the whisking frequency averaged out across trials (Fig. 3f). The average of L6-evoked movements is described by the outline ('envelope') of one or several whisking bursts (Fig. 3f). The power spectrum of the average movement revealed that only a single, low-frequency (<0.5 Hz) peak was present at frequencies that were contained in the envelope of a whisking burst. The same observations were made for the population average ($n = 11$; Supplementary Fig. 1b). Averaged movements evoked by L5 and L6 cell stimulation were significantly different in peak power frequency (unpaired t -test, $P < 0.001$), whereas the peak power frequencies of individual trials were not (unpaired t -test, $P > 0.2$; Supplementary Fig. 1c). Thus, across trials the whisking rhythm is stable when L5 cells are stimulated, but averages out when L6 cells are stimulated.

Effects of AP frequency and number

To explore how AP frequency in M1 pyramidal cells influences movements, we studied whisker movements evoked by trains of 10 APs, initiated at rates of 10, 50 or 100 Hz. In an L6 pyramidal cell (Fig. 4a), 10 APs at 50 or 100 Hz evoked a sequence of bursts of

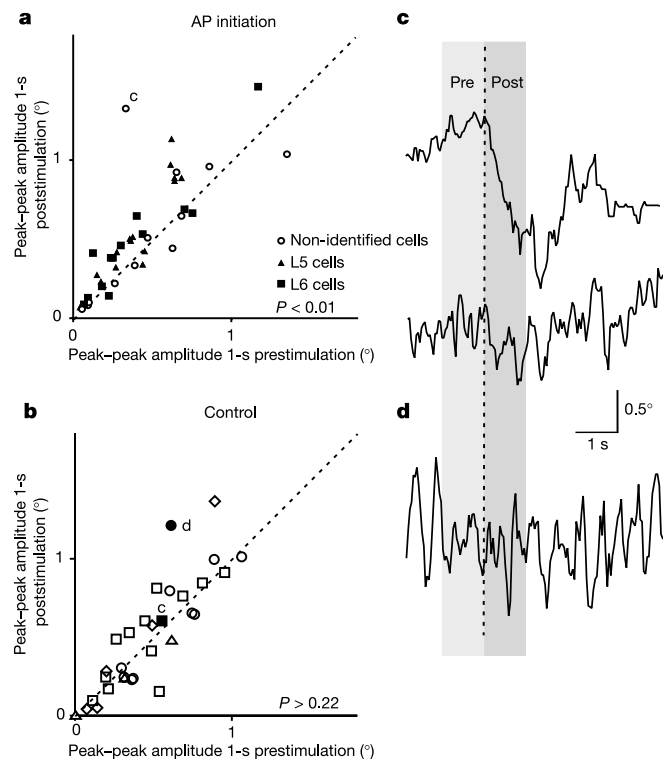


Figure 2 Action potential initiation specifically evokes whisker movements. **a**, Pre- and poststimulation amplitudes of whisker movement during single-cell stimulation in 44 neurons. Data are based on averaged movements (10–30 trials) evoked by stimulation with 10 APs at either 50 Hz (26 cells) or 100 Hz (18 cells). **b**, Pre- and poststimulation amplitudes of whisker movement in control trials. Circles, extracellular injection of 5-nA current steps; squares, intracellular injection of hyperpolarizing current steps; triangles, intracellular injection of depolarizing but subthreshold current steps; diamonds, APs in neurons of primary visual cortex. **c**, Movement trace of the largest evoked movement (top) and the respective control trace with hyperpolarizing current injection (bottom). **d**, Largest (random) movement effect observed under control conditions. Broken lines in **c** and **d** indicate the onset of stimulation. The time intervals used for determining the movement amplitudes shown in **a** and **b** are indicated by grey shading in **c** and **d**.

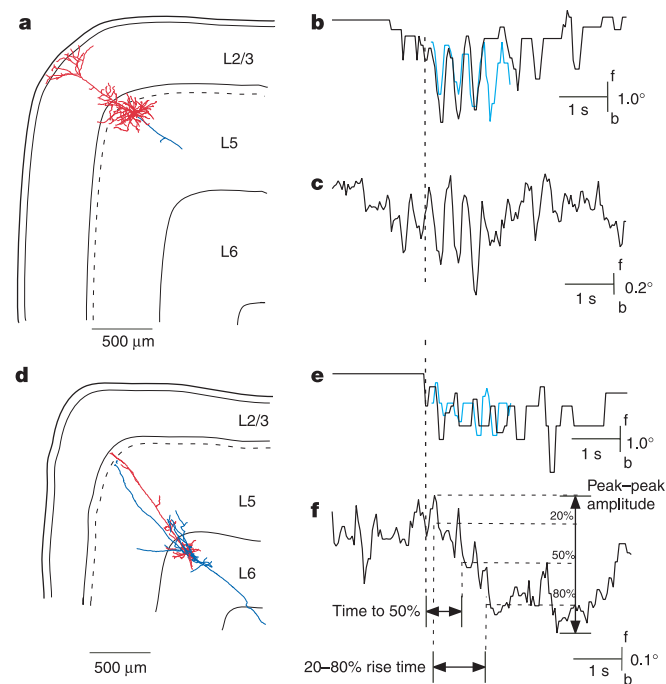


Figure 3 Whisker movements evoked by intracellular stimulation of single L5 and L6 cells. **a**, Dendrites (red) and axon (blue) of the stimulated L5 neuron. Proximal extracellular stimulation evoked backward movement of a single whisker (D1). **b**, Position of whisker D1 in a partially (blue) and fully (black) shown intracellular stimulation trial (10 APs at 50 Hz). Note that the evoked movements are in phase. **c**, Movement average (15 trials). **d**, Dendrites (red) and axon (blue) of the stimulated L6 neuron. Proximal extracellular stimulation evoked backwards movements of several whisker rows, with maximal deflection of whisker C1. **e**, Position of whisker C1 in a partially (blue) and fully (black) shown intracellular stimulation trial (10 APs at 50 Hz). Note that the evoked movements are not in phase. **f**, Movement average (15 trials). Movement parameters were determined as indicated. Broken lines in **b**, **c**, **e** and **f** indicate the onset of AP initiation.

backwards whisking movements. In the averaged movement (Fig. 4b), the outlines (envelopes) of two such bursts are apparent: the amplitudes of the evoked backwards movements were substantially larger at 100-Hz than at 50-Hz stimulation. Action potentials

at 10 Hz resulted in a marked, slow, forwards movement after the train. Six additional experiments gave the same result (Fig. 4c, d and Supplementary Fig. 2a).

The effect of AP frequency on the peak-to-peak amplitude was significant (analysis of variance (ANOVA), $F_{(2,6)} = 12.18$; $P < 0.003$), whereas effects on the 20–80% rise time (ANOVA, $F_{(2,6)} = 0.9$; $P > 0.4$) and the time to 50% amplitude (ANOVA, $F_{(2,6)} = 0.34$; $P > 0.7$) were insignificant. The effects of variation of AP frequency were similar for L5 and L6 cells. Consistent with the results from intracellular stimulation, we observed a reversal of whisker movement direction at frequencies of less than 50 Hz for extracellular stimulation (data not shown).

To investigate the effect of AP number, we carried out experiments in which 2, 5 or 10 APs were evoked at a rate of 50 Hz in an L6 cell (Fig. 4e, f). With an increasing number of APs, the latency and rise time of the evoked backwards movements decreased

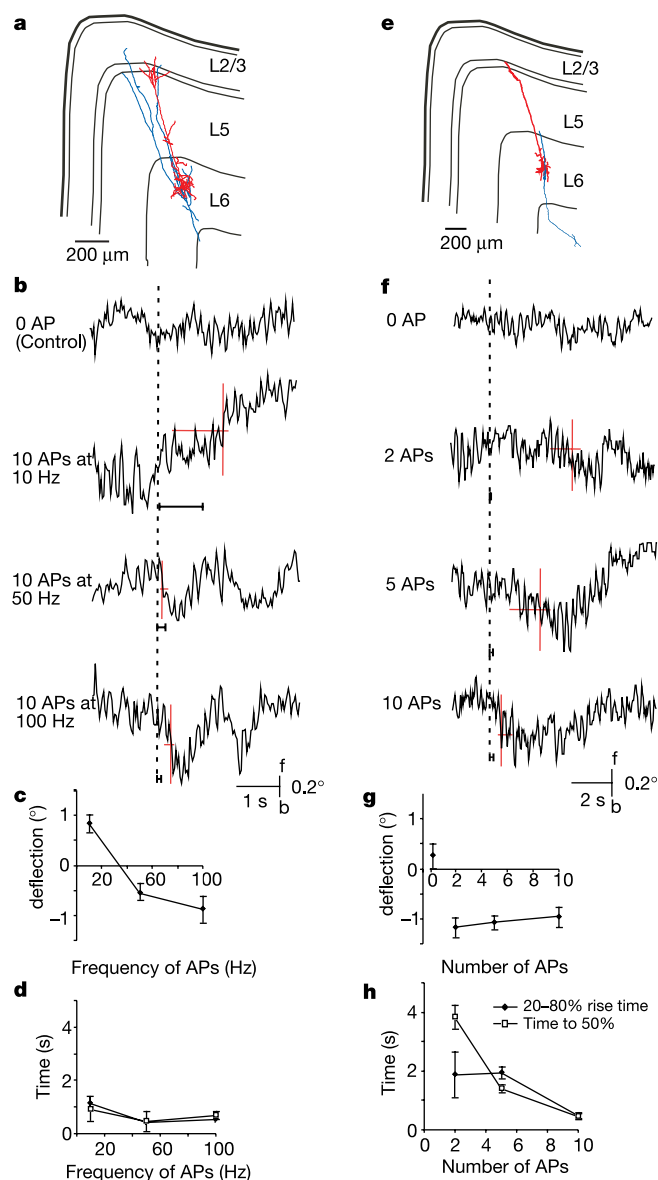


Figure 4 Effect of AP number and frequency on evoked movements. **a**, Dendrites (red) and axon (blue) of the stimulated L6 neuron. Proximal extracellular stimulation evoked backwards movements of C-row whiskers, with maximal deflection of whisker C2. **b**, Average movements of whisker C2 in 15 stimulation trials with initiation of 10 APs at 10, 50 or 100 Hz. Vertical red lines indicate the 50% amplitude time point and the amplitude measured; horizontal red lines indicate the 20–80% rise times. **c**, Population data for the effect of AP frequency on mean deflection amplitudes. Data are from seven neurons (three L5 pyramidal cells, three L6 pyramidal cells and one unidentified regularly spiking cell) located at sites where extracellular stimulation evoked backwards movements. **d**, Effect of AP frequency on movement dynamics (same cells as in **c**). **e**, Dendrites (red) and axon (blue) of the stimulated L6 neuron. Proximal extracellular stimulation evoked backwards movements of posterior whiskers in rows C, D and E, with maximal deflection of whisker E1. **f**, Average movements of whisker E1 in 12 intracellular stimulation trials with initiation of 2, 5 or 10 APs at 50 Hz. Red lines are the same as in **b**. **g**, Population data for the effect of AP number on mean deflection amplitudes. Data are from nine neurons (two L5 pyramidal cells, four L6 pyramidal cells and three unidentified regularly spiking cells) located at sites where extracellular stimulation evoked backwards movements. **h**, Effect of AP number on movement dynamics (same cells as in **g**). Broken lines in **b** and **f** indicate the onset of AP initiation.

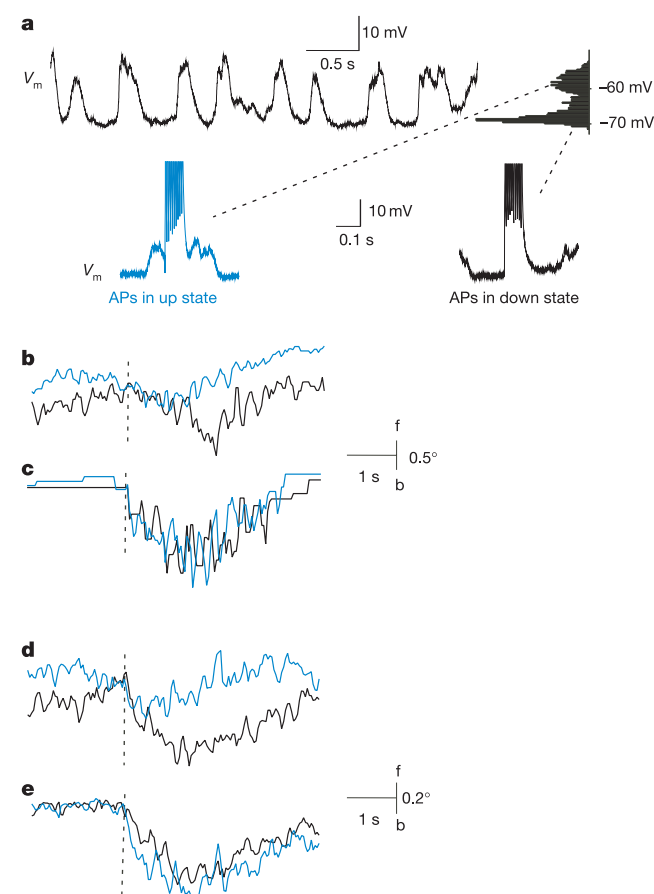


Figure 5 Efficacy of APs initiated in cortical up states and down states with and without prestimulation movement. **a**, Top, the membrane potential of an L6 cell (same cell as in Fig. 4e) fluctuated between more depolarized (up states) and less depolarized (down states), leading to a bimodal distribution of membrane potential values over time (histogram on right). Bottom, AP initiation (10 APs at 100 Hz) in up states (left, blue) or down states (right, black); APs have been clipped. **b**, Average evoked movements for APs initiated in up states (blue, $n = 8$ trials) or down states (black, $n = 12$ trials). **c**, Average evoked movements with AP initiation in down states (black, $n = 3$ trials) or up states (blue, $n = 6$ trials) for stimulation trials without prestimulation movements. **d**, Population average (133 trials, $n = 9$ cells in which large backwards movements were evoked (three L5 pyramidal cells, three L6 pyramidal cells and three unidentified regularly spiking cells)) of movements evoked by AP initiation in down states (black trace) or up states (blue trace). **e**, Population average (56 trials, $n = 9$ cells) of movements evoked by AP initiation in down states (black trace) or up states (blue trace), when trials with prestimulation movements were excluded. Broken lines in **b–e** indicate the onset of AP initiation.

progressively (Fig. 4f). The amplitude remained either constant or increased slightly. Eight additional experiments gave the same result (Fig. 4g, h and Supplementary Fig. 2b). The time course of movements evoked by fewer than 10 APs was very slow. Because the motor effects of initiating 1 AP (data not shown) and 2 APs were variable from cell to cell and from trial to trial, further work will be required to determine the robustness of the effects of single APs on movements.

The effect of AP number (for 1, 2, 5 or 10 APs) on the time to 50% amplitude was more pronounced than the effect of AP number on the 20–80% rise time (Fig. 4h), indicating that AP number affects latency more than rise time. An increasing AP number significantly decreased the 20–80% rise time (ANOVA, $F_{(3,8)} = 9.18$; $P < 0.02$) and the time to 50% amplitude (ANOVA, $F_{(3,8)} = 8.18$; $P < 0.01$), but had no significant effect on the peak-to-peak amplitude (ANOVA, $F_{(3,8)} = 0.42$; $P > 0.4$). The increase of movement latency with decreasing AP number was observed in both L5 and L6 cells.

Interaction of APs with ongoing cortical activity

Under the anaesthesia used in this study, the membrane potential of M1 neurons fluctuates by 10–15 mV between depolarized (up states) and hyperpolarized (down states) membrane potentials (Fig. 5a). Up states and down states reflect synchronous large-amplitude fluctuations of cortical population activity^{27–29}. To understand how the motor effects of single-cell discharges depend on the ongoing cortical activity, we sorted data into stimulation trials in which APs were initiated either in up states (Fig. 5a, bottom, blue trace) or down states (Fig. 5a, bottom, black trace). Down-state AP initiation evoked larger movements (1.3°; Fig. 5b, black trace) than did up-state AP initiation (0.5°; Fig. 5b, blue trace). The pattern of cortical up states and down states was closely related to the ongoing whisker movements of the rat, with whisker movements occurring preferentially in prolonged up states. Thus,

movements evoked by APs in the up state sometimes counteracted the trajectory of ongoing movements.

To minimize the effects of ‘spontaneous’ movements associated with up states and down states, we carried out an analysis in which we excluded stimulation trials with spontaneous (prestimulation) movement (Fig. 5c). For this subset of data, down-state AP initiation (2.0°; Fig. 5c, black trace) and up-state AP initiation (1.7°; Fig. 5c, blue trace) led to very similar evoked movements. At the population level, AP initiation in up states (0.35°; Fig. 5d, blue trace) evoked smaller movements than did AP initiation in down states (1.05°; Fig. 5d, black trace), but, in the absence of prestimulation movements, AP efficacy in up and down states was similar (1.31° versus 1.15°; Fig. 5e).

Thus, APs initiated in up states are seemingly less effective only because of the effects of ongoing movement patterns associated with cortical up and down states. This view was supported when we corrected evoked movements by subtracting the average movement patterns associated with up states and down states (Supplementary Fig. 3). Thus, much of the variability in the single-cell stimulation effects might be attributed to fluctuations in ongoing cortical population activity and associated movements.

AP trains are more effective in awake rats

We observed progressively larger evoked whisker movements with a decreasing depth of anaesthesia (data not shown; see also refs 7, 26, 30, 31). This effect was studied more closely in four cells that we stimulated under conditions of both anaesthesia and wakefulness. In one L6 cell (Fig. 6a), no movement was evoked by intracellular stimulation (10 APs at 100 Hz) under anaesthesia (Fig. 6b), but a very large whisker movement could be evoked by the same stimulation in the awake rat (Fig. 6c). In this restricted sample, whisker movements were evoked by AP initiation in the awake condition (mean ratio of post- to prestimulation amplitudes, 2.2 ± 0.7), but not under the anaesthetized condition (ratio of

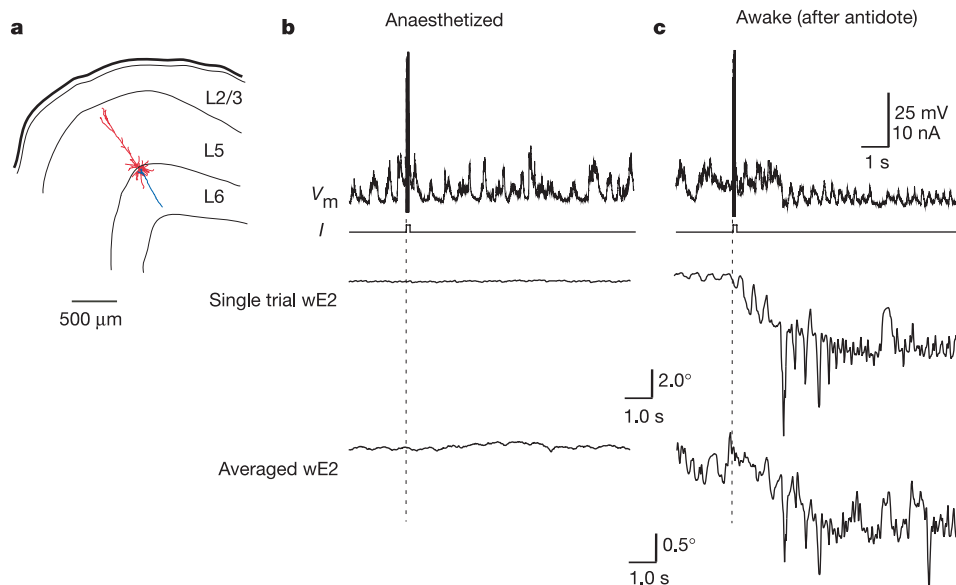


Figure 6 Cortical APs are more effective in awake rats. **a**, Dendrites (red) and axon (blue) of the stimulated L6 neuron. Proximal extracellular stimulation evoked backwards movement of several whisker rows, with maximal deflection of whisker E2 (wE2). **b**, Intracellular stimulation in a head-fixed rat anaesthetized with a mixture of medetomidine, midazolam and fentanyl. Top to bottom, the four traces show membrane potential, the stimulation current, position of whisker E2 during a single stimulation (10 APs at 100 Hz) trial, and average position of whisker E2 for ten stimulation trials. **c**, Stimulation of the same cell after anaesthesia was antagonized with a mixture of

atipamezol, flumazenil and naloxone. A local field potential recording indicated a marked change in the pattern of population activity (not shown). The top three traces indicate the last stimulation trial, the bottom trace indicates the averaged movement trace of the last ten stimulation trials (~4 min after antagonizing anaesthesia) before the recording was disrupted owing to a rapid body movement of the rat. Seconds later, head fixation was discontinued and the rat began to explore the recording set-up. Conventions are the same as in **b**.

post- to prestimulation amplitudes, 0.94 ± 0.08).

For the four cells stimulated in awake rats, we observed a larger evoked slow whisker movement (Fig. 6c) and a larger evoked fast whisker movement (data not shown) than in any of the 44 cells stimulated in anaesthetized rats. These data, albeit restricted to a small sample, indicate that anaesthesia dampens the movement effects and that AP initiation in the vibrissae motor cortex of awake rats is highly effective.

Discussion

We have shown that a train of APs in a single pyramidal cell of rat M1 can cause whisker movement, thereby contradicting the view that “the discharge of an individual pyramidal neuron in M1 is meaningless”³². Three results support our conclusion: first, currents applied intracellularly were about four orders of magnitude smaller (nanoamperes versus tens of microamperes) than for extracellular stimulation, making it unlikely that unintended stimulation through an extracellular pathway contaminated the results; second, the patterning of single-cell-evoked movements was dependent on the layer in which the stimulated cell was located; and third, evoked movements varied with the number and frequency of initiated AP trains.

We observed complex and prolonged movement patterns that involved repeated movement of several whiskers. Thus, in the vibrissae motor cortex, individual pyramidal cells seem to specify motor programs rather than single muscle contractions. It seems that the lamination of M1 mirrors specific properties of whisking, such as the timing of the individual whisking movement by L5 activity and the grouping of multiple whisking movements in bursts of whisking by L6 activity. Extracellular single-unit recordings have described AP activity in deep-layer M1 cells that was correlated with the initiation of whisker movements but not with the phase of whisking³³, consistent with our observations on single-cell stimulation effects in L6. A laminar difference in stimulation effects has been also indicated by extracellular stimulation experiments in cats³⁴, in which stimulation of deeper layers evoked more prolonged and less phasic contractions of limb muscles than did more superficial stimulation.

Movements evoked by intracellular stimulation occur with long latency and greatly outlast the duration of the AP train. The direction of movements is dependent on the stimulation frequency. These observations are indicative of indirect control of whisker movements by neurons of the vibrissae motor cortex. Indeed, cells in M1 do not directly project to the facial nucleus, which innervates the whisker musculature³⁵, but reach the facial nucleus through disynaptic pathways. A principal pathway consists of a projection from M1 L5 cells to an ensemble of cells distributed around the facial nucleus, which are thought to form a ‘central pattern generator’ for whisking movements^{35–37}. The complexity of the evoked movements might result from activation of the putative pattern generator. Our data indicate that L5 cells do not merely gate an otherwise freely running pattern generator. Instead, APs in single L5 cells can either reset or initiate the whisking rhythm. L6 neurons project to the thalamus³⁸, and the bursts of whisking movements after L6 cell stimulation might indicate that thalamic circuits are involved in the grouping of individual protraction and/or retractions into bursts of whisking.

An intriguing issue is how a small number of APs evoked in a single neuron can be so effective in the extensive M1 network. It might be that our whole-cell recordings tended to sample a subset of neurons—for example, particularly large neurons—that might be unusually effective in evoking movements. However, we have no evidence for such a sampling bias. Initiation of multiple APs might recruit both intracortical and subcortical postsynaptic target neurons by temporal summation³⁹ and by synaptic facilitation, which is a prominent feature of corticofugal output synapses⁴⁰. On the basis of its surface area, we estimate that the agranular medial area, from

which whisker movements can be evoked, contains 1–1.5 million neurons⁴¹. The largest whisker movements made by rats have amplitudes of around 100° , whereas single-cell stimulation in anaesthetized rats often evoked multiwhisker movements with an amplitude of around 0.5° . Thus, it seems possible that the activity of only a few hundred neurons could be sufficient to achieve the full set of whisker movements.

One fact that might contribute to the effectiveness of single-cell activity is the low rate of spontaneous APs in cortical neurons^{42–45}. Movements evoked by single cells seem to vary as a result of variations in ongoing cortical activity and the associated movements. Taking this source of variance into account, our data indicate that there is an unanticipated degree of precision and sensitivity in the mechanisms that ‘translate’ APs into movements. We conclude that, in vibrissae motor cortex, sparse cortical activity is sufficient to generate motor output. □

Methods

Rats

We used standard surgical and electrophysiological techniques⁴⁵. Wistar rats ($n = 63$) were anaesthetized with ketamine hydrochloride (90 mg per kg (body weight) intraperitoneally) and xylazine (5 mg per kg intraperitoneally) for surgery. Pressure points and skin incisions were infused with lidocaine, and rats received supplemental doses of ketamine (20 mg per kg) and acepromazine (0.02 mg per kg, intramuscularly) as needed. Vibrissae motor cortex (A1–A2.5 and L1–L2; see Supplementary Information) and, in control trials, visual cortex (P5 and L3.5) were exposed. All experimental procedures were carried out in accordance with the animal welfare guidelines of the Max Planck Society.

Stimulation and recordings

Extracellular stimulation (300-ms trains of 100 monophasic, negative, 0.3-ms stimulation pulses) was delivered through a pipette (tip resistance $0.5\text{--}1\text{ M}\Omega$) at $1,500\text{ }\mu\text{m}$ below the pial surface. Intracellular stimulation consisted of 5-ms or 10-ms current steps for evoking individual APs in sequences, as specified in the text. Intracellular stimulation was delivered at a rate of 0.1 or 0.2 Hz for 10–30 trials per condition. We obtained recordings as described⁴⁵ and filled the pipettes with (in mM) 130 potassium gluconate, 10 sodium gluconate, 10 HEPES, 10 phosphocreatine, 4 Mg-ATP, 2 Na₂-ATP, 0.3 GTP, 4 NaCl and 0.4% biocytin (pH 7.2).

All recordings were done in cortical L5 and L6. We identified L5 by the large size of cell somata. In 70% of our whole-cell recordings of pyramidal cells, where the resting membrane potential was stable and no holding currents were applied, the ongoing AP activity was very low (mean 0.21 ± 0.6 Hz, median 0 Hz). We excluded 6 out of 50 recordings from M1 neurons from analysis because the anaesthesia was considered to be too deep (the criterion was an absence of even, small whisker movements; that is, fewer than one whisker movement of $>0.5^\circ$ per 20 s). Thus, 44 stimulation experiments with 34 regularly spiking and 10 intrinsically bursting cells were analysed.

We made four additional whole-cell recordings to compare anaesthetized and awake conditions. For these, rats (P28–P30; $n = 6$) were prepared for chronic recordings⁴⁵ and habituated to head fixation. On the day of the experiment, rats were anaesthetized with an intraperitoneal dose of a mixture of Dormitor (medetomidine hydrochloride (Roche); 225 μg per kg), Dormicum (midazolam (Roche); 6 μg per kg) and fentanyl (Janssen-Cilag) (7.5 μg per kg). A whole-cell recording was established, and anaesthesia was antagonized with a subcutaneous dose of a mixture of Antisedan (atipamezole (Roche); 1 mg per kg), Anexate (flumazenil (Roche); 600 μg per kg) and Nalcanti (naloxone (Bristol Myers-Squibb); 180 μg per kg). Antagonizing anaesthesia required 1–3 min.

Whisker movement analysis

We videotaped whisker movements from directly above, and the angular movements of whiskers of interest were quantified in a frame-by-frame analysis of videotapes. A subset of the data was analysed with specially developed tracking software (Whiskerwatcher, Arrington Research); the precision of the tracking system usually exceeded 0.05° . Whisker position data refer to the ‘best whisker’ (the maximally displaced whisker) determined by extracellular stimulation. We report peak-to-peak poststimulus movement amplitudes (measured as shown in Fig. 3f) rather than the baseline-to-peak amplitudes. Measurement of peak-to-peak poststimulus movement amplitudes and baseline-to-peak amplitudes led to similar results except for 1-AP and 2-AP initiation, where evoked movements often consisted of a small forwards movement followed by a large and slow backwards movement. We also analysed, trial by trial, the root mean square amplitude of whisker movements. This analysis led to quantitative conclusions very similar to the quantification of peak-to-peak amplitudes (data not shown). Movements were sampled at a rate of 25 Hz (in two experiments at both 25 Hz and 150 Hz), which is sufficient to sample the rather slow whisker movements of anaesthetized rats.

Histology

Cells were recovered by histological procedures⁴⁴ and biocytin-labelled neurons were reconstructed with NeuroLucida software (MicroBrightfield). Data are reported as the mean $\pm$ s.e.m. unless otherwise stated.

Received 1 July; accepted 3 December 2003; doi:10.1038/nature02266.

1. Fritsch, G. & Hitzig, E. Über die elektrische Erregbarkeit des Grosshirns. *Arch. Anat. Physiol. Wiss. Med.* **37**, 300–332 (1870).
2. Asanuma, H. *The Motor Cortex* (Raven, New York, 1989).
3. Porter, R. & Lemon, R. *Corticospinal Function and Voluntary Movement* (Clarendon, Oxford, 1995).
4. Asanuma, H. & Sakata, H. Functional organization of a cortical efferent system examined with focal depth stimulation in cats. *J. Neurophysiol.* **30**, 35–54 (1967).
5. Adrian, E. D. & Moruzzi, G. Impulses in the pyramidal tract. *J. Physiol. (Lond.)* **100**, 159–191 (1939).
6. Evarts, E. V. Activity of pyramidal tract neurons during postural fixation. *J. Neurophysiol.* **32**, 375–385 (1969).
7. Stoney, S. D. Jr, Thompson, W. D. & Asanuma, H. Excitation of pyramidal tract cells by intracortical microstimulation: effective extent of stimulating current. *J. Neurophysiol.* **31**, 659–669 (1968).
8. Asanuma, H., Stoney, S. D. Jr & Abzug, C. Relationship between afferent input and motor outflow in cat motorsensory cortex. *J. Neurophysiol.* **31**, 670–681 (1968).
9. Georgopoulos, A. P. Current issues in directional motor control. *Trends Neurosci.* **18**, 506–510 (1995).
10. Schwartz, A. B. & Moran, D. W. Arm trajectory and representation of movement processing in motor cortical activity. *Eur. J. Neurosci.* **6**, 1851–1856 (2000).
11. Chapin, J. K., Moxon, K. A., Markowitz, R. S. & Nicolelis, M. A. Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. *Nature Neurosci.* **7**, 664–670 (1999).
12. Wessberg, J. *et al.* Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature* **408**, 361–365 (2000).
13. Taylor, D. M., Tillery, S. I. & Schwartz, A. B. Direct cortical control of 3D neuroprosthetic devices. *Science* **296**, 1829–1832 (2002).
14. Schwartz, A. B., Taylor, D. M. & Tillery, S. I. Extraction algorithms for cortical control of arm prosthetics. *Curr. Opin. Neurobiol.* **11**, 701–707 (2001).
15. Fetz, E. E., Cheney, P. D. & German, D. C. Corticomotoneuronal connections of precentral cells detected by postspike averages of EMG activity in behaving monkeys. *Brain Res.* **114**, 505–510 (1976).
16. Cheney, P. D. & Fetz, E. E. Comparable patterns of muscle facilitation evoked by individual corticomotoneuronal (CM) cells and by single intracortical microstimuli in primates: evidence for functional groups of CM cells. *J. Neurophysiol.* **53**, 786–804 (1985).
17. Buys, E. J., Lemon, R. N., Mantel, G. W. H. & Muir, R. B. Selective facilitation of different hand muscles in the conscious monkey. *J. Physiol. (Lond.)* **381**, 529–549 (1986).
18. Woody, C. D. & Black-Cleworth, P. Differences in excitability of cortical neurons as a function of motor projection in conditioned cats. *J. Neurophysiol.* **36**, 1104–1116 (1973).
19. Lerma, J. & Garcia-Austt, E. Hippocampal theta rhythm during paradoxical sleep. Effects of afferent stimuli and phase relationships with phasic events. *Electroencephalogr. Clin. Neurophysiol.* **60**, 46–54 (1985).
20. Timo-Iaria, C., Yamashita, R., Hoshino, K. & Souza-Melo, A. Rostrum movements in desynchronized sleep as a prevalent manifestation of dreaming activity in Wistar rats. *Braz. J. Med. Biol. Res.* **23**, 617–620 (1990).
21. Bermejo, R., Harvey, M., Gao, P. & Zeigler, H. P. Conditioned whisking in the rat. *Somatosens. Mot. Res.* **13**, 225–233 (1996).
22. Bermejo, R., Vyas, A. & Zeigler, H. P. Optoelectronic monitoring of whisking trajectories in two dimensions. The topography of the rat's 'whisking space'. *Soc. Neurosci. Abstr.* **51.8** [online] (2001).
23. Carvell, G. E. & Simons, D. J. Biometric analyses of vibrissal tactile discrimination in the rat. *J. Neurosci.* **10**, 2638–2648 (1990).
24. Welker, W. I. Analysis of sniffing of the albino rat. *Behaviour* **22**, 223–244 (1964).
25. Kleinfeld, D., Berg, R. W. & O'Connor, S. M. Anatomical loops and their electrical dynamics in relation to whisking by rat. *Somatosens. Mot. Res.* **16**, 69–88 (1999).
26. Hall, R. D. & Lindholm, E. P. Organization of motor and somatosensory neocortex in the albino rat. *Brain Res.* **66**, 23–38 (1974).
27. Anderson, J., Lampl, I., Reichova, I., Carandini, M. & Ferster, D. Stimulus dependence of two-state fluctuations of membrane potential in cat visual cortex. *Nature Neurosci.* **3**, 617–621 (2000).
28. Cowan, R. L. & Wilson, C. J. Firing patterns and axonal projections of single corticostriatal neurons in the rat medial agranular cortex. *J. Neurophysiol.* **71**, 17–32 (1994).
29. Petersen, C. C., Hahn, T. T., Mehta, M., Grinvald, A. & Sakmann, B. Interaction of sensory responses with spontaneous depolarization in layer 2/3 barrel cortex. *Proc. Natl Acad. Sci. USA* **100**, 13638–13643 (2003).
30. Leyton, A. S. F. & Sherrington, C. S. Observations on the excitable cortex of the chimpanzee, orang-utan and gorilla. *Q. J. Exp. Physiol.* **11**, 135–222 (1917).
31. Berg, R. W. & Kleinfeld, D. Vibrissa movement elicited by rhythmic electrical microstimulation to motor cortex in the aroused rat mimics exploratory whisking. *J. Neurophysiol.* **90**, 2950–2963 (2003).
32. Eccles, J. C. *Understanding of the Brain* 2nd edn (McGraw Hill, New York, 1976).
33. Carvell, G. E., Miller, S. A. & Simons, D. J. The relationship of vibrissal motor cortex unit activity to whisking in the awake rat. *Somatosens. Mot. Res.* **13**, 115–127 (1996).
34. Asanuma, H. & Ward, J. E. Patterns of contraction of distal forelimb muscles produced by intracortical stimulation in cats. *Brain Res.* **27**, 97–109 (1971).
35. Hattori, A. M., Priest, C. A. & Keller, A. Functional circuitry involved in the regulation of whisker movements. *J. Comp. Neurol.* **442**, 266–276 (2002).
36. Gao, P., Bermejo, R. & Zeigler, H. P. Whisker deafferentation and rodent whisking patterns: behavioral evidence for a central pattern generator. *J. Neurosci.* **21**, 5374–5380 (2001).
37. Hattori, A., Li, Y. & Keller, A. Serotonin regulates rhythmic whisking. *Neuron* **17**, 343–352 (2003).
38. Jones, E. G. in *Cerebral Cortex* Vol. 1 (eds Jones, E. G. & Peters, A.) 521–553 (Plenum, New York, 1984).
39. Jankowska, E., Padel, Y. & Tanaka, R. The mode of activation of pyramidal tract cells by intracortical tract cells. *J. Physiol. (Lond.)* **249**, 617–636 (1975).
40. Phillips, C. G. & Porter, R. *Corticospinal Neurons. Their Role in Movement* (Academic, London, 1977).
41. Rockel, A. J., Hiorns, R. W. & Powell, T. P. The basic uniformity in structure of the neocortex. *Brain* **103**, 221–244 (1980).
42. Moore, C. I. & Nelson, S. B. Spatio-temporal subthreshold receptive fields in the vibrissa representation of rat primary somatosensory cortex. *J. Neurophysiol.* **80**, 2882–2892 (1998).
43. Zhu, J. J. & Connors, B. W. Intrinsic firing patterns and whisker-evoked synaptic responses of neurons in the rat barrel cortex. *J. Neurophysiol.* **81**, 1171–1183 (1999).
44. Brecht, M. & Sakmann, B. Dynamic representation of whisker deflection by postsynaptic potentials in morphologically reconstructed spiny stellate and pyramidal cells in the barrels and septa of layer 4 in rat somatosensory cortex. *J. Physiol. (Lond.)* **543**, 49–70 (2002).
45. Margrie, T. W., Brecht, M. & Sakmann, B. *In vivo*, low-resistance, whole-cell recordings from neurons in the anaesthetized and awake mammalian brain. *Pflügers Arch. Eur. J. Physiol.* **444**, 491–498 (2002).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank I. Manns, R. Friedrich, C. Schwarz, W. Denk and S. Petrou for comments and R. Erickson for inspiration; E. Heil, M. Kaiser, R. Rödel, P. Mayer and K. Schmidt for technical assistance; and A. Krauss, S. Muhammad, S. Bellanca and L. Sinai-Esfahani for help with cell staining and reconstruction. This work was supported by the Max Planck Society, the NHMRC of Australia and the Wellcome Trust.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.B. (brecht@mpimf-heidelberg.mpg.de).

Decoherence of matter waves by thermal emission of radiation

Lucia Hackermüller, Klaus Hornberger, Björn Brezger*, Anton Zeilinger & Markus Arndt

¹Institut für Experimentalphysik, Universität Wien, Boltzmanngasse 5, A-1090 Wien, Austria

* Present address: Fachbereich Physik, Universität Konstanz, D-78457 Konstanz, Germany

Emergent quantum technologies have led to increasing interest in decoherence—the processes that limit the appearance of quantum effects and turn them into classical phenomena. One important cause of decoherence is the interaction of a quantum system with its environment, which ‘entangles’ the two and distributes the quantum coherence over so many degrees of freedom as to render it unobservable. Decoherence theory^{1–4} has been complemented by experiments using matter waves coupled to external photons^{5–7} or molecules⁸, and by investigations using coherent photon states⁹, trapped ions¹⁰ and electron interferometers^{11,12}. Large molecules are particularly suitable for the investigation of the quantum–classical transition because they can store much energy in numerous internal degrees of freedom; the internal energy can be converted into thermal radiation and thus induce decoherence. Here we report matter wave interferometer experiments in which C₇₀ molecules lose their quantum behaviour by thermal emission of radiation. We find good quantitative agreement between our experimental observations and microscopic decoherence theory. Decoherence by emission of thermal radiation is a general mechanism that should be relevant to all macroscopic bodies.

In this Letter we investigate the decoherence of molecular matter waves. We change the internal temperature of the molecules in a controlled way before they enter a near-field interferometer, and observe the corresponding reduction of the interference contrast. The idea behind this effort is to demonstrate a most fundamental decoherence mechanism that we encounter in the macroscopic world: all large objects, but also molecules of sufficient complexity, are able to store energy and to interact with their environment via thermal emission of photons. It is generally believed that warm macroscopic bodies emit far too many photons to allow the observation of de Broglie interferences, whereas individual atoms

or molecules can be sufficiently well isolated to exhibit their quantum nature. However, there must be a transition region between these two limiting cases. Interestingly, as we show in this study, C₇₀ fullerene molecules have just the right amount of complexity to exhibit perfect quantum interference in our experiments¹³ at temperatures below 1,000 K, and to gradually lose all their quantum behaviour when the internal temperature is increased up to 3,000 K. We can thus trace the quantum-to-classical transition in a controlled and quantitative way. The complexity of large molecules adds a novel quality with respect to previously performed experiments with atoms^{5–7}: the energy in molecules may be equilibrated in many internal degrees of freedom during the free flight, and a fraction of the vibrational energy will eventually be reconverted into emitted photons. Therefore the internal dynamics of the molecule is also relevant for the quantum behaviour of the centre-of-mass state. In contrast to resonance fluorescence, which was investigated with atoms^{5–7}, thermal decoherence is omnipresent in macroscopic systems and it cannot be switched off.

The basic set-up of our experiment¹⁴ is sketched in Fig. 1: a beam of C₇₀ molecules is generated by sublimation at about 900 K. The molecules pass a heating stage where they cross a focused argon ion laser beam up to 16 times. The fullerenes interact with the laser approximately every 0.3 mm. The laser heating increases the molecular temperature by 140 K per absorbed photon. We calculate that they reach up to 5,000 K for very short times, but the re-emission of thermal photons is so efficient that even the hottest molecules are cooled to below ~3,000 K when they enter the interferometer 7.2 cm behind the heating stage.

The interferometer consists of three identical free-standing gold gratings with a period of $d = 991$ nm. They are separated by the equal distance of $L = 38$ cm, which is the Talbot length $L_T = d^2/\lambda_{dB}$ for a typical de Broglie wavelength of $\lambda_{dB} = 2.6$ pm. The first grating acts as a periodic array of narrow slit sources, the second one as the diffracting element, and the third grating is used as a scanning detection mask, which modulates the molecular density pattern produced by the Talbot–Lau interference effect^{15,16}. The transmitted molecules are ionized by a blue laser beam (wavelength 488 nm, 6.6 μ m waist), and their intensity I is recorded as a function of the lateral displacement of the third grating. The fringe visibility $V = (I_{max} - I_{min})/(I_{max} + I_{min})$ characterizes the interferogram and thereby the coherence of the molecular evolution.

The essence of the experiment is to measure the variation of the interference fringe visibility with heating laser power (Fig. 2). Two observations can be made: first, the interference contrast decreases

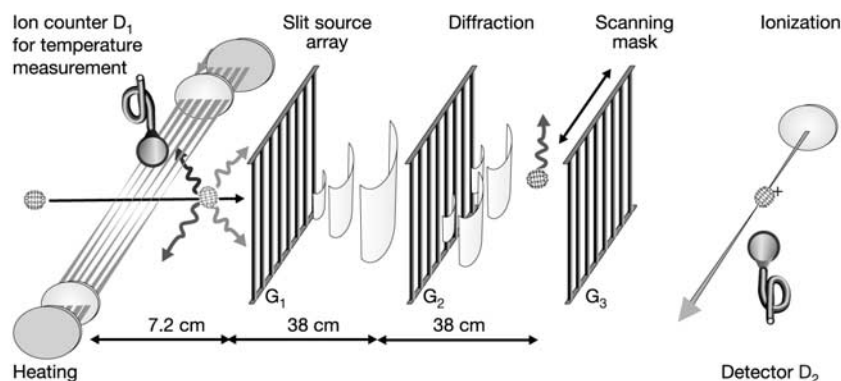


Figure 1 Set-up for the observation of thermal decoherence in a Talbot–Lau molecule interferometer. A fullerene beam passes from left to right, interacting with a heating stage, a three-grating (G₁–G₃) matter-wave interferometer and an ionizing detection laser beam in D₂ (wavelength 488 nm, 1/e² intensity radius 6.6 μ m, 15 W). The gold gratings have a period of 991 nm and slit widths of nominally 475 ± 20 nm. Decoherence of the fullerene

matter waves can be induced by heating the molecules with multiple laser beams (514.5 nm, 40 μ m waist radius, 0–10 W) before they enter the interferometer. The resulting molecular temperature can be assessed by detecting the heating-dependent fraction of fullerene ions using the electron multiplier D₁ over the heating stage.

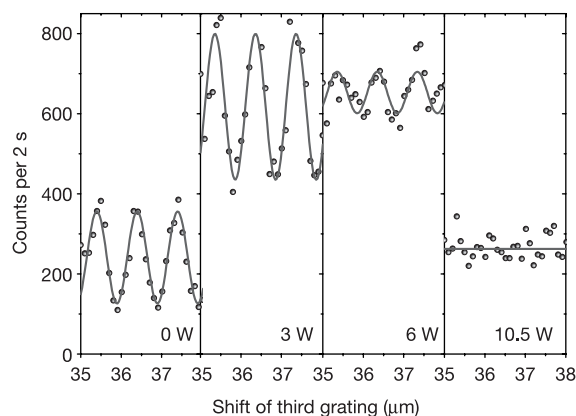


Figure 2 Molecule interferograms for C_{70} at 190 m s^{-1} for increasing laser heating powers, P . The fringe visibility V decreases with increasing heating power P owing to the rising emission probability of visible photons: $P = 0 \text{ W}$ ($V = 47\%$), $P = 3 \text{ W}$ ($V = 29\%$), $P = 6 \text{ W}$ ($V = 7\%$), $P = 10.5 \text{ W}$ ($V = 0\%$). In contrast to that, the absolute count rate grows initially with increasing P . This is due to the fact that the thermal ionization probability in detector D_2 increases with the temperature of the arriving molecules. At even higher heating intensities the count rate falls again because of ionization and fragmentation in the heating stage.

monotonically with increasing power, and vanishes at 10 W. This is the signature of decoherence due to the enhanced probability for the emission of thermal photons that carry ‘which-path’ information. Second, we notice that the count rate also varies considerably. This is explained by the dependence of the ionization efficiency in the detector D_2 on the internal energy of the fullerenes. It proves that much internal energy remains in the molecules during their flight through the apparatus.

In order to confirm quantitatively the interpretations of both observations, we model the evolution of the distribution of the internal energies on their way through the apparatus. The temperature dependence of the spectral photon emission rate (equation (1) below) then yields the loss of fringe visibility as predicted by decoherence theory (equation (2) below).

The first photon absorption populates the electronic triplet state T_1 via the excited singlet S_1 . Given the known C_{70} triplet lifetimes and non-radiative transition rates (see ref. 17 and references therein), we can assume that all further excitation occurs in the triplet system and that the absorbed excess energy is rapidly transferred to the vibrational levels. It is known that fullerenes may store more than 100 eV for a very short time¹⁷, and it was observed that at high temperatures three different cooling mechanisms start to compete—the thermal emission of photons, electrons or C_2 dimers^{18–22}. These processes are the molecular analogues of the bulk phenomena known as blackbody radiation, thermionic emission and evaporative cooling. Following the most recent experimental data²², we may safely assume that fragmentation is the least efficient mechanism. In contrast, thermally activated ionization is an important mechanism, which we use both in our fullerene detector²³ and for molecule thermometry, as discussed below. Nevertheless, we can safely neglect both delayed ionization and fragmentation for the discussion of the fringe contrast, because the recoil upon fragmentation and ionization is generally so large that the affected molecules will miss the narrow detector. We have also experimentally confirmed that neither C_{70}^+ ions nor C_{68} nor smaller fragments from the heating region are recorded by the detector D_2 .

However, C_{70}^+ ions—and potentially ionized fragments—can be detected immediately above the heating stage by the electron multiplier D_1 (Fig. 1). To get an estimate of the molecular temperature distribution, we record the number of ions as a function of

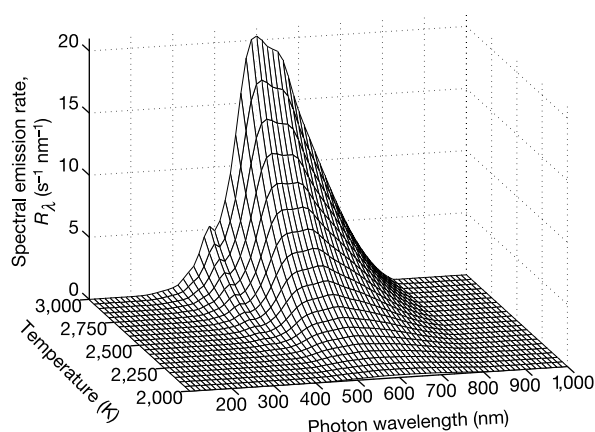


Figure 3 Spectral photon emission rate R_λ of C_{70} molecules, as used for the calculation of thermal decoherence. We use the published²⁵ absorption cross-section for ($S_0 \rightarrow S_1$) and a heat capacity of $C_V = 202 k_B$. The fall-off to short wavelengths is determined by the limited internal energy of the molecules, while the decrease at long wavelengths is due to the lack of accessible radiative transitions at energies below $\sim 1.5 \text{ eV}$. The figure shows that in the absence of cooling a single molecule at 2,500 K travelling at 190 m s^{-1} (that is, with a transit time of 4 ms through the interferometer) would emit an integrated number of three visible photons. This is sufficient to determine the path of the molecule if the emission occurs close to the second grating.

the heating power and of the fullerene velocity. By comparing the data with a model calculation, we can extract the parameters that govern the molecular heating of C_{70} . Our model describes the spatial and velocity-dependent distribution of the internal molecular energy by accounting for the stochastic absorption process, the laser beam characteristics, and the rapid radiative cooling between the beams as determined by equation (1) below. It reproduces the detected number of ions in the heating stage for different laser powers, different numbers of heating beams and all velocities with the fit parameters for the triplet absorption cross-section, $\sigma(T_1) = 2 \times 10^{-17} \text{ cm}^2$, and the effective Arrhenius constant for ionization, $A_{\text{ion}} = 5 \times 10^9 \text{ s}^{-1}$. The same calculation also describes the heating-dependent increase in count rate at the detector D_2 and thus yields independent information on the temperature distribution in the molecular beam.

The mean temperature in the beam drops rapidly behind the heating stage through the emission of thermal photons. The emission of a continuous photon spectrum has already been observed for fullerenes in other experiments^{18,24}. The equation for the thermal radiation density differs from the macroscopic Planck law for several reasons. First, the thermal wavelengths are much larger than the size of the fullerene, turning it into a coloured emitter. The mean emission probability is proportional to the usual mode density factor $\omega^2/\pi^2 c^2$ and the known frequency-dependent absorption cross-section²⁵ $\sigma_{\text{abs}}(\omega)$, assuming that it does not strongly depend on the internal temperature. Second, the particle is not in thermal equilibrium with the radiation field. It emits into a cold environment and stimulated emission does not occur. For this reason, the statistical factor $1/[\exp(\hbar\omega/k_B T) - 1]$ of the Planck formula now would read $\exp(-\hbar\omega/k_B T)$. Third, the 204 vibrational modes of C_{70} do not constitute an infinite heat bath but have a finite heat capacity C_V . Therefore the emission does not take place at a fixed temperature, although the internal energy is nonetheless conveniently characterized by the micro-canonical temperature T_m . This leads to a further correction in the spectral photon emission rate²⁶, which is now fully described by

$$R_\omega(\omega, T_m) = \frac{\omega^2}{\pi^2 c^2} \sigma_{\text{abs}}(\omega) \exp \left[-\frac{\hbar\omega}{k_B T_m} - \frac{k_B}{2C_V} \left(\frac{\hbar\omega}{k_B T_m} \right)^2 \right] \quad (1)$$

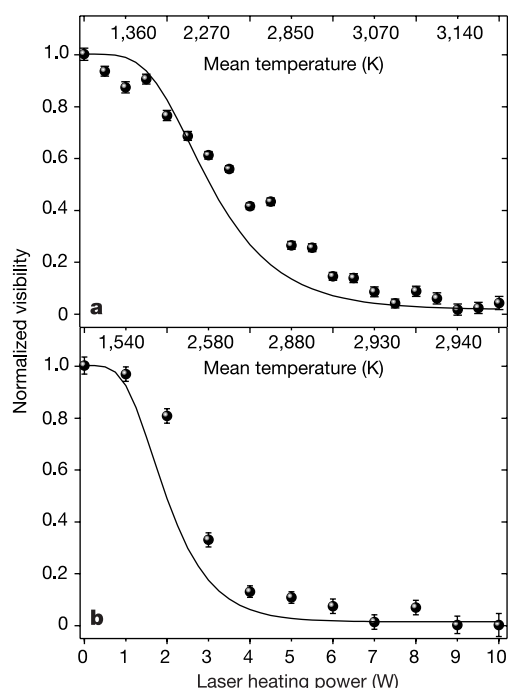


Figure 4 Decoherence curves. **a**, Interference visibility as a function of laser heating power (lower scale). The molecular beam with a mean velocity of $v_m = 190 \text{ m s}^{-1}$ passes a $50 \mu\text{m}$ central height delimiter comparable to the waist ($40 \mu\text{m}$) of the 16 heating laser beams. We observe a rapid decrease of the fringe visibility with increasing power both in the experiment (circles) and in theory (solid line). The upper axis indicates the mean temperature of the molecules when they enter the interferometer. The maximum contrast without heating was $V_0 = 47\%$, which is close to the theoretical value¹¹. **b**, Molecules with $v_m = 100 \text{ m s}^{-1}$, selected by a $150 \mu\text{m}$ height delimiter and heated by ten beams of the specified incident laser power. The qualitative behaviour is the same and the quantitative agreement with theory is as good as before. The maximum contrast for this velocity class was $V_0 = 19\%$. In both experimental arrangements, a mean number between one and two visible photons is required to reduce the contrast by a factor of two.

In Fig. 3 we plot the wavelength dependence of $R_\lambda = R_\omega |d\omega/d\lambda|$. We observe that at temperatures below $2,000 \text{ K}$ the emission rate is negligible, whereas at higher temperatures the molecules may emit photons whose wavelengths are comparable to (or even smaller than) the maximum path separation of $\sim 1 \mu\text{m}$. They transmit (partial) which-path information to the environment, leading to a reduced observability of the fullerene wave nature. Around $3,000 \text{ K}$ the molecules have a high probability to emit several visible photons yielding sufficient which-path information to effect a complete loss of fringe visibility in our interferometer.

A formal description of this qualitative picture can be given by decoherence theory. It considers the entanglement of the molecule with the emitted photon, and shows how coherences vanish once a trace over the photon state is performed. For objects with velocity v and temperature evolution $T(t)$ we obtain a visibility

$$V = V_0 \exp \left[- \int_0^{2L/v} dt \int_0^\infty d\lambda R_\lambda(\lambda, T(t)) \times \left\{ 1 - \text{sinc} \left(2\pi \frac{d}{\lambda} \cdot \frac{L - |vt - L|}{L_T} \right) \right\} \right] \quad (2)$$

as discussed in the Methods section. V_0 denotes the interference contrast in the absence of photon emission. In the exponential, the sinc function compares the effective molecular path separation

to the radiation wavelength, while the integrals cover all photon wavelengths λ and longitudinal positions vt in the interferometer. As a result, the visibility is reduced exponentially whenever photons are emitted whose wavelength is sufficiently small to resolve the path separation. Our predictions for the loss of visibility are obtained by weighting equation (2) with the previously determined distribution of temperature evolutions in the molecular beam.

In Fig. 4 we compare our decoherence model with the experiments by plotting the interference fringe visibility as a function of the laser power. We observe a strong decrease of the visibility for molecules at 190 m s^{-1} , heated by 16 laser beams (Fig. 4a), and for molecules at 100 m s^{-1} , heated by 10 laser beams (Fig. 4b).

We also observe good agreement between decoherence theory (solid line) and the experiment (circles). The experiment is reproducible within the indicated error bars for a given laser alignment, but small displacements of the laser focus will influence the shape and slope of the observed decoherence curve. The difference between the theoretical and the experimental curve is of the order of this variation.

In summary, we have presented conclusive empirical and numerical evidence for observation of the quantum-to-classical transition of a material object caused by its own emission of thermal radiation. This auto-localization is a fundamental process limiting the ultimate observability of quantum effects in macroscopic objects. However, for nanometre-sized systems^{13,27,28} this mechanism becomes relevant only at high temperatures, and it is not expected to be a limitation for interference of objects even considerably larger than the fullerenes, such as proteins. $\square$

Methods

Equation (2) describes the loss of matter wave coherence due to the emission of thermal photons. It is obtained by assuming that the emission is isotropic and that the absorbing walls of the apparatus are located in the far-field, where the photon position distribution reflects its momentum distribution. In this case, a trace over the photon state changes the fullerene centre-of-mass state $\hat{\rho}$ according to

$$\hat{\rho} \rightarrow \hat{\rho}' = \int d\mathbf{k} \frac{p(k)}{4\pi k^2} \hat{U}_k \hat{\rho} \hat{U}_k^\dagger \quad (3)$$

where the $\hat{U}_k = \exp(i\mathbf{r}\mathbf{k})$ are momentum translation operators and $p(k)$ is the probability density for the photon wavenumber $k = 2\pi/\lambda$. In the position representation of the density matrix,

$$\rho'(\mathbf{r}_1, \mathbf{r}_2) = \langle \mathbf{r}_1 | \hat{\rho}' | \mathbf{r}_2 \rangle = \langle \mathbf{r}_1 | \hat{\rho} | \mathbf{r}_2 \rangle \eta(\mathbf{r}_1 - \mathbf{r}_2) \quad (4)$$

we find from equation (3) that the off-diagonal elements are reduced by the decoherence function²⁹

$$\eta(\mathbf{r}_1 - \mathbf{r}_2) = \frac{1}{R_{\text{tot}}} \int_0^\infty d\lambda R_\lambda(\lambda) \text{sinc} \left(2\pi \frac{|\mathbf{r}_1 - \mathbf{r}_2|}{\lambda} \right) \quad (5)$$

Here $p(k)$ is expressed in terms of the spectral emission rate R_λ (see equation (1)) and the total photon emission rate R_{tot} . This sinc-shaped position dependence of η is also found in other experiments with isotropic momentum change^{7,8}. It describes the diffraction limitation of a hypothetical microscope used to obtain which-path information on the molecules.

In the Talbot-Lau geometry^{14,16,27} the final molecular fringe pattern $w(x)$ is strictly periodic in the grating constant d , and can be expanded as a Fourier series, which reads in the absence of decoherence

$$w(x) = \sum_l C_l \exp(2\pi i l x / d) \quad (6)$$

Assuming that a single photon emission occurs at the longitudinal position $z = vt$, a closed expression for the resulting molecular density pattern can be found. It is obtained by propagating the molecular density matrix in paraxial approximation first to the position z . We then apply the decohering transformation (equation (4)) followed by a propagation to the final grating. For a set-up with equally spaced and identical gratings, the new fringe pattern is described by a simple modification of the Fourier coefficients

$$C_l \rightarrow C_l' = C_l \cdot \eta(l d (L - |vt - L|) / L_T) \quad (7)$$

In order to account for more than one photon emission, we make the Markov assumption that all photon emissions are independent of each other. Because the modification

(equation (7)) is independent of the molecular density matrix, the change of the final density pattern is governed by the differential equation

$$\frac{d}{dt} C_i = R_{\text{tot}} \left[C_i \eta \left(\frac{L - |vt - L|}{L_T} \right) - C_i \right] \quad (8)$$

It describes how the final interferogram is blurred as the time interval of emission increases. Equation (2) follows then immediately after taking into account the time dependence of the emission rate due to cooling (equation (1)) and the fact that for our grating geometry, with a slit width of 470 nm and grating constant of 990 nm, only the lowest-order Fourier components contribute to the fringe visibility³⁰ $V = 2|C_1/C_0|$.

Received 15 September; accepted 8 December 2003; doi:10.1038/nature02276.

- Joos, E. *et al.* *Decoherence and the Appearance of a Classical World in Quantum Theory* (Springer, Heidelberg, 2003).
- Blanchard, P., Giulini, D., Joos, E., Kiefer, C. & Stamatescu, I.-O. (eds) *Decoherence: Theoretical, Experimental, and Conceptual Problems* (Lecture Notes in Physics 538, Springer, Heidelberg, 2000).
- Imry, Y. *Introduction to Mesoscopic Physics*, 2nd edn (Oxford Univ. Press, Oxford, 2000).
- Zurek, W. H. Decoherence, einselection, and the quantum origins of the classical. *Rev. Mod. Phys.* **75**, 715–775 (2003).
- Pfau, T., Spälter, S., Kurtsiefer, Ch., Ekstrom, C. R. & Mlynek, J. Loss of spatial coherence by a single spontaneous emission. *Phys. Rev. Lett.* **73**, 1223–1226 (1994).
- Chapman, M. S. *et al.* Photon scattering from atoms in an atom interferometer: coherence lost and regained. *Phys. Rev. Lett.* **75**, 3783–3787 (1995).
- Kokorowski, D. A., Cronin, A. D., Roberts, T. D. & Pritchard, D. E. From single- to multiple-photon decoherence in an atom interferometer. *Phys. Rev. Lett.* **86**, 2191–2194 (2001).
- Hornberger, K. *et al.* Collisional decoherence observed in matter wave interferometry. *Phys. Rev. Lett.* **90**, 160401 (2003).
- Brune, M. *et al.* Observing the progressive decoherence of the “meter” in a quantum measurement. *Phys. Rev. Lett.* **77**, 4887–4890 (1996).
- Myatt, C. J. *et al.* Decoherence of quantum superpositions through coupling to engineered reservoirs. *Nature* **403**, 269–273 (2000).
- Buks, E., Schuster, R., Heiblum, M., Mahalu, D. & Umansky, V. Dephasing in electron interference by a ‘which-path’ detector. *Nature* **391**, 871–874 (1998).
- Ji, Y. *et al.* An electronic Mach–Zehnder interferometer. *Nature* **422**, 415–418 (2003).
- Arndt, M. *et al.* Wave–particle duality of C_{60} molecules. *Nature* **401**, 680–682 (1999).
- Brezger, B. *et al.* Matter-wave interferometer for large molecules. *Phys. Rev. Lett.* **88**, 100404 (2002).
- Patorski, K. in *Progress in Optics XXVII* (ed. Wolf, E.) 2–108 (Elsevier Science, Amsterdam, 1989).
- Clauser, J. F. & Li, S. Talbot-von Lau atom interferometry with cold slow potassium. *Phys. Rev. A* **49**, R2213–R2217 (1994).
- Dresselhaus, M. S., Dresselhaus, G. & Eklund, P. C. *Science of Fullerenes and Carbon Nanotubes* (Academic, San Diego, 1996).
- Mitzner, R. & Campbell, E. E. B. Optical emission studies of laser desorbed C_{60} . *J. Chem. Phys.* **103**, 2445–2453 (1995).
- Ding, D., Huang, J., Compton, R. N., Klotz, C. E. & Haufler, R. E. Cw laser ionization of C_{60} and C_{70} . *Phys. Rev. Lett.* **73**, 1084–1087 (1994).
- Hansen, K. & Echt, O. Thermionic emission and fragmentation of C_{60} . *Phys. Rev. E* **78**, 2337–2340 (1997).
- Kolodney, E., Tsipinyuk, B. & Budrevich, A. The thermal energy dependence (10–20 eV) of electron impact induced fragmentation of C_{60} in molecular beams: experiment and model calculations. *J. Chem. Phys.* **102**, 9263–9275 (1995).
- Matt, S. *et al.* Kinetic energy release distribution and evaporation energies for metastable fullerene ions. *Chem. Phys. Lett.* **303**, 379–386 (1999).
- Nairz, O., Arndt, M. & Zeilinger, A. Experimental challenges in fullerene interferometry. *J. Mod. Opt.* **47**, 2811–2821 (2001).
- Hesler, P., Carlsson, J. O. & Demirev, P. Photon emission from gas phase fullerenes excited by 193 nm laser radiation. *Chem. Phys.* **107**, 10440–10445 (1997).
- Coheur, P. F., Carleer, M. & Colin, R. The absorption cross sections of C_{60} and C_{70} in the visible-UV region. *J. Phys. B* **29**, 4987–4995 (1996).
- Hansen, K. & Campbell, E. E. B. Thermal radiation from small particles. *Phys. Rev. E* **58**, 5477–5482 (1998).
- Hackermüller, L. *et al.* Wave nature of biomolecules and fluorofullerenes. *Phys. Rev. Lett.* **91**, 090408 (2003).
- Clauser, J. F. in *Experimental Metaphysics* (eds Cohen, R. S., Home, M. & Stachel, J.) 1–11 (Kluwer Academic, Dordrecht, 1997).
- Joos, E. & Zeh, H. D. The emergence of classical properties through interaction with the environment. *Z. Phys. B* **59**, 223–243 (1985).
- Brezger, B., Arndt, M. & Zeilinger, A. Concepts for near-field interferometers with large molecules. *J. Opt. B* **5**, 82–89 (2003).

Acknowledgements We thank S. Uttenthaler for his support in an early stage of the experiment. We acknowledge support by the Austrian START programme, the Austrian FWF, the European TMR and Marie Curie programmes, and the DFG Emmy-Noether programme.

Authors' contributions L.H. performed most of the experiments as a part of her Ph.D. thesis. K.H. developed the decoherence theory, and made the quantitative comparison between experiment and theory.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.A. (markus.arndt@univie.ac.at).

High-transition-temperature superconductivity in the absence of the magnetic-resonance mode

J. Hwang¹, T. Timusk¹ & G. D. Gu²

¹Department of Physics and Astronomy, McMaster University, Hamilton, Ontario L8S 4M1, Canada

²Department of Physics, Brookhaven National Laboratory, Upton, New York 11973-5000, USA

The fundamental mechanism that gives rise to high-transition-temperature (high- T_c) superconductivity in the copper oxide materials has been debated since the discovery of the phenomenon. Recent work has focused on a sharp ‘kink’ in the kinetic energy spectra of the electrons as a possible signature of the force that creates the superconducting state^{1–14}. The kink has been related to a magnetic resonance^{13,15–17} and also to phonons¹⁸. Here we report that infrared spectra of $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ (Bi-2212), shows that this sharp feature can be separated from a broad background and, interestingly, weakens with doping before disappearing completely at a critical doping level of 0.23 holes per copper atom. Superconductivity is still strong in terms of the transition temperature at this doping ($T_c \approx 55$ K), so our results rule out both the magnetic resonance peak and phonons as the principal cause of high- T_c superconductivity. The broad background, on the other hand, is a universal property of the copper–oxygen plane and provides a good candidate signature of the ‘glue’ that binds the electrons.

We investigated the Bi-2212 material systematically as a function of doping, including highly overdoped Bi-2212 (see Methods). To better display the sharp ‘kink’, we do not show the optical conductivity but focus our attention on a related quantity, the optical single-particle self-energy, $\Sigma^{\text{op}}(\omega)$ (see Methods). In Fig. 1a–d we plot the imaginary part of this quantity, which is simply the scattering rate of the charge carriers. At high frequencies the scattering rate varies linearly with frequency: this is called the marginal Fermi-liquid (MFL) behaviour¹⁹. We note that the overall scattering rate decreases as the doping increases and that there is a sharp depression in $1/\tau(\omega)$ below 700 cm^{-1} at low temperatures, with an overshoot in the superconducting state. This sharp onset of scattering has been in general attributed to the interaction of the charge carriers with a bosonic mode and more recently to the 41-meV neutron resonance^{7,13,14}. In the real part of the optical self-energy, plotted in Fig. 1e–h, we note a sharp peak around 700 cm^{-1} , which tracks the depressions in $1/\tau(\omega)$ in both frequency and amplitude. One interesting property of the peak in the self-energy is its doping and temperature dependence: the peak gets weaker as the doping level and the temperature increase. We call the peak the ‘optical resonance mode’. The resonance peak in the self-energy spectrum can clearly be resolved from the broader background of MFL scattering.

The optical self-energy is closely related to the quasiparticle self-energy $\Sigma^{\text{qp}}(\omega)$, which can be measured in ARPES experiments although there are important differences in the two quantities^{20–22}. In Fig. 2 we show a comparison of the optical self-energy from our data and the ARPES self-energy at two temperatures: one in the normal state and the other in the superconducting state. We see that although the qualitative features are very similar, the magnitudes and the frequency values of the various features differ considerably, as expected from theory. Part of the difference can be attributed to the uncertainty in locating the unrenormalized baseline in the ARPES analysis. As noted in ref. 6, a peak in $\Sigma^{\text{qp}}(\omega)$ can be clearly separated from the broad continuum and is closely correlated with

the neutron resonance mode. The lower noise level and the higher resolution of our optical data demonstrate a clear doping-dependent trend—the amplitude of the resonance peak weakens with doping in the overdoped region and disappears completely at a doping level where the superconductivity is still large enough to show $T_c \approx 55$ K, as shown in Fig. 2g. Thus, at this doping level we have superconductivity without the resonance peak, and whereas the peak may contribute to the condensation energy^{16,23} and the superconducting gap at lower doping levels, it cannot be the main cause of high- T_c superconductivity.

In Fig. 3 we investigate further the disappearance of the optical resonance mode as a function of doping. In Fig. 3a we plot the amplitude of the resonance mode in the real part of the optical self-energy, which decreases uniformly with doping and appears to extrapolate to zero at doping level $p = 0.23$ ($T_c \approx 55$ K). The centre frequency of the mode (Fig. 3b) is proportional to the transition temperature ($\Omega_{\text{res}}^{\text{op}} \approx 8.0k_B T_c$) and reaches a maximum at the optimally doped phase for both FTIR and ARPES, which is consistent with other studies⁴. In Fig. 3c we show the contribution of the resonance mode to the imaginary part of the self-energy as a function of doping, both directly and normalized as a percentage of the total scattering rate at $3,000 \text{ cm}^{-1}$ (see Methods). This is done to divide out the overall decrease of the coupling of the charge carriers to the bosonic fluctuations that occurs with increased doping^{6,24}.

We do not have detailed data on the amplitude of the mode as a

function of temperature. In the region of $p = 0.19$ the mode appears at the superconducting transition temperature of 80 K along with the neutron mode^{3,4}. The error in the critical doping value cannot be found from optical data alone because we were not able to obtain a sample with sufficient overdoping. The ARPES data of ref. 6 for their $T_c = 55$ K sample shows no resonance, so we can use this as an upper limit and our own 60-K sample as a lower limit, giving $p = 0.225 \pm 0.010$ for the value of the critical doping point.

Although it is clear from our results that the amplitude of the coupling of the optical resonance mode to charge carriers extrapolates to zero at a critical doping of $p = 0.23$, we are not able to determine the behaviour of the centre frequency of the resonance mode at the critical doping point. The centre frequency appears to drop below the parabolic trend, shown as the dashed curve in Fig. 3b. According to theory, the frequency of the ARPES resonance ($\Omega_{\text{res}}^{\text{ap}}$) is roughly the sum of the gap (Δ) and the frequency of the neutron mode¹³ (Ω) (we note that in optics $\Omega_{\text{res}}^{\text{op}} = 2\Delta + \Omega$).

We remark here on the relationship between our work and the proposal of Lanzara *et al.*¹⁸ that the sharp kink seen in angle-resolved photoemission is due to phonons. With our higher resolution and lower noise level we are able to separate the sharp peak from the overall broad background. The broad background

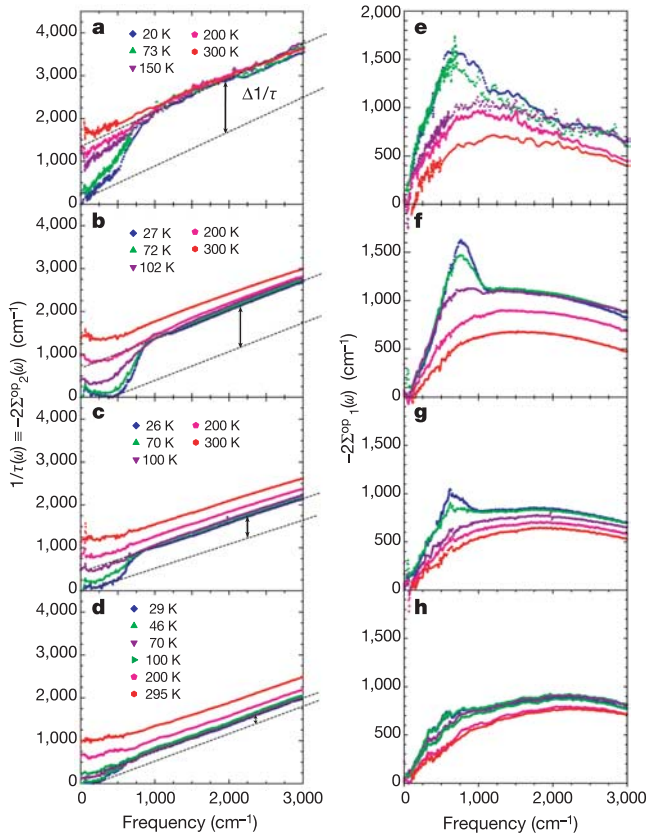


Figure 1 The optical single-particle self-energy of $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$. **a–d**, The doping and temperature dependent optical scattering rate, $1/\tau(\omega)$ for four representative doping levels. **a**, $T_c = 67$ K (underdoped); **b**, 96 K (optimal); **c**, 82 K (overdoped); **d**, 60 K (overdoped). **e–h**, The real part of the optical self-energy, $-2\Sigma_1^{\text{op}}(\omega)$. The slope of $-\Sigma_1^{\text{op}}(\omega)$ near $\omega = 0$ is closely related to the coupling constant or the mass enhancement factor, and also decreases as the doping increases. This is consistent with other studies^{6,24}. We note the weakening of the feature at 700 cm^{-1} in both sets of curves as the doping level increases.

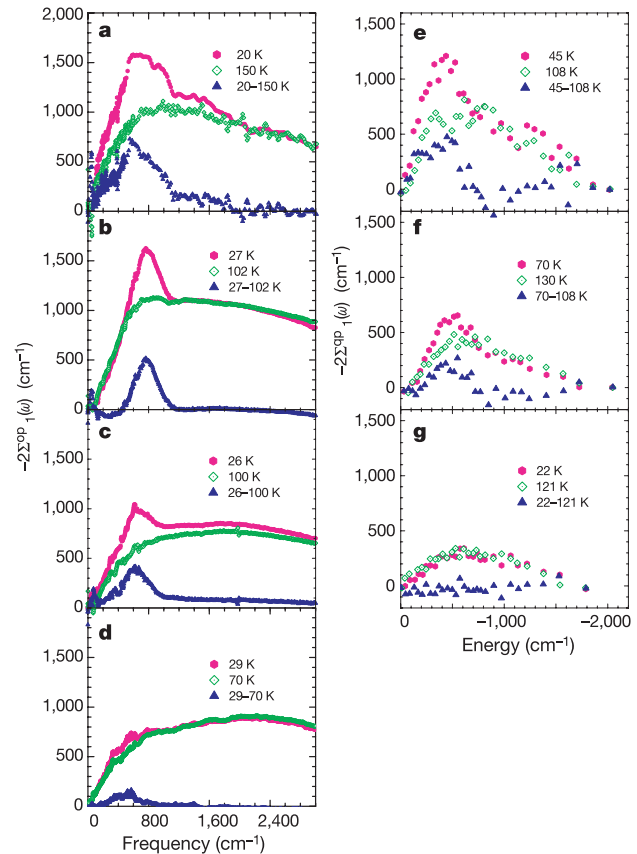


Figure 2 Comparison of the self-energy measured with infrared and angle-resolved photoemission for Bi-2212. **a–d**, The real part of the optical self-energies, $-2\Sigma_1^{\text{op}}(\omega)$ of the normal and superconducting phase curves. **a**, $T_c = 67$ K (underdoped); **b**, $T_c = 96$ K (optimal); **c**, $T_c = 82$ K (overdoped); **d**, $T_c = 60$ K (overdoped). **e–g**, The real part of the self-energies, $-2\Sigma_1^{\text{ap}}(\omega)$ from the ARPES measurements of ref. 6. **e**, $T_c = 69$ K (underdoped); **f**, $T_c = 91$ K (optimal); **g**, $T_c = 55$ K (overdoped). Although the absolute magnitudes and frequencies differ for the two data sets, the overall qualitative features are closely correlated when one allows for the higher noise level and lower resolution of the photoemission data.

can be seen in the spectra of all high- T_c superconductors and seems to represent a universal property of the copper oxide plane. But it cannot be due to phonons, because its spectral weight extends beyond the cut-off frequency of the phonon spectrum²². The sharp structure, on the other hand, has all the characteristics of magnetic resonance: it makes its appearance at the superconducting T_c in overdoped copper oxides and slightly above T_c in the underdoped copper oxides. We have also shown that it vanishes completely at a high doping level inside the superconducting dome. None of this is expected for phonons, which should be present at all temperatures and doping levels.

The disappearance of the resonance mode may be related to the possible vanishing of the pseudogap and the existence of a quantum critical point^{25,26}. Recent data²⁷ suggest that the critical doping level for the disappearance of the pseudogap in c -axis transport may be higher than $p = 0.22$, at least in Bi-2212. The pseudogap has less influence than the resonance mode on a - b -plane properties, as shown by work on the anisotropy of d.c. transport²⁸. Also on theoretical grounds, a gap in the fluctuation spectrum will not give the upward step in the scattering rate that we observe⁷. The disappearance of the resonance mode has been predicted to occur near the $p = 0.22$ doping level within the magnetic exciton picture^{9,10}. Our analysis of the optical self-energy, which allows us to isolate the resonance peak from the universal broad background in the single-particle self-energy for Bi-2212, should work for other copper oxides and gives us a direct tool for examining the

spectrum of excitations responsible for pairing in high- T_c superconductors. □

Methods

Experiments and sample preparations

We used Fourier-transform infrared (FTIR) spectroscopy to obtain the reflectance data²⁹ of floating-zone-grown single-twinned crystals of Bi-2212. The highly overdoped samples were made by annealing the as-grown crystals at high temperature, $T > 500^\circ\text{C}$, in 3,000-bar liquid oxygen in sealed containers. We present new data on one optimally doped Bi-2212 ($T_c = 96\text{ K}$), and three highly overdoped samples ($T_c = 82\text{ K}$, $T_c = 65\text{ K}$, and $T_c = 60\text{ K}$). The four new samples are bulk superconductors with London penetration depths of 2,065 Å, 1,638 Å, 1,481 Å and 1,375 Å, respectively. The optimally doped sample has been doped with a small amount of Y to yield a relatively well-ordered system³⁰. To cover a wide range of doping levels we also used data from previously measured underdoped Bi-2212 crystals¹².

The optical single-particle self-energy

The optical conductivity was determined by Kramers–Kronig analysis²⁹ and we present our data within the framework of the extended Drude model, where the complex optical conductivity can be written as $\sigma(\omega) = -i\omega\epsilon(\omega) - \epsilon_H/4\pi = -i\omega_p^2/[4\pi(2\Sigma^{\text{op}}(\omega) - \omega)]$, where ϵ_H is the dielectric constant at a high frequency ($\sim 2\text{ eV}$) for each doping level (note that there is doping dependence in ϵ_H), ω_p is the plasma frequency and $\Sigma^{\text{op}}(\omega) = \Sigma^{\text{op}}(\omega) + i\Sigma_2^{\text{op}}(\omega)$ is the optical single-particle self-energy. In terms of more familiar quantities $\Sigma_1^{\text{op}}(\omega) \equiv \omega(1 - m^*/m)/2$, where m^*/m is the mass renormalization, and $\Sigma_2^{\text{op}}(\omega) \equiv -1/(2\tau(\omega))$, where τ is the carrier lifetime. We determine the doping-dependent plasma frequency from the absorption spectra in the near-infrared region³⁴.

Doping dependence of the optical resonance mode in $\Sigma^{\text{op}}(\omega)$

We used the parabolic approximation to obtain the doping levels from the transition temperature, T_c , with the slope of the infrared reflectance as an additional test of the doping level³⁴. The amplitude and centre frequency of the optical resonance mode are obtained from $-2\Sigma_2^{\text{op}}(\omega)$ data by a least-squares fit of the peak to a lorentzian line shape (see Fig. 3a and b). To estimate the strength of the interaction of the resonance mode with the carriers from $1/\tau(\omega) \equiv -2\Sigma_2^{\text{op}}(\omega)$, we draw a dashed line, parallel to the high-frequency trend, from the onset point of substantial scattering. The difference between this line and the actual high-frequency scattering, $\Delta 1/\tau$ is our estimate of the contribution of the sharp mode to the scattering. It exhibits a doping dependence that is suggestive of mean-field behaviour (dashed curve) with a critical point at $p = 0.23$ (see Fig. 1a–d and Fig. 3c). We estimate the doping dependence of the onset temperature of the resonance mode from both FTIR and ARPES data. We have drawn the points as follows: the upper error-limit temperature is where there is no resonance. The lower-limit temperature is where we first see the resonance. The resonance onset temperature is the average of the upper and lower limits (see Fig. 3a).

Received 12 August 2003; accepted 16 January 2004; doi:10.1038/nature02347.

- Rossat-Mignod, J. *et al.* Neutron scattering study of the $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ system. *Physica C* **185**, 86–92 (1991).
- Mook, H. A. *et al.* Spin fluctuations in $\text{YBa}_2\text{Cu}_3\text{O}_{6.6}$. *Nature* **395**, 580–582 (1998).
- Fong, H. F. *et al.* Neutron scattering from magnetic excitations in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$. *Nature* **398**, 588–591 (1999).
- He, H. *et al.* Resonant spin excitation in an overdoped high temperature superconductor. *Phys. Rev. Lett.* **86**, 1610–1613 (2001).
- Kaminski, A. *et al.* Renormalization of spectral line shape and dispersion below T_c in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$. *Phys. Rev. Lett.* **86**, 1070–1073 (2001).
- Johnson, P. D. *et al.* Doping and temperature dependence of the mass enhancement observed in the cuprate $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$. *Phys. Rev. Lett.* **87**, 177007 (2001).
- Norman, M. R. & Ding, H. Collective modes and the superconducting-state spectral function of $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+x}$. *Phys. Rev. B* **57**, R11089–R11092 (1998).
- Campuzano, J. C. *et al.* Electronic spectra and their relation to the (π, π) collective mode in high- T_c superconductors. *Phys. Rev. Lett.* **83**, 3709–3712 (1999).
- Abanov, A. *et al.* A relation between the resonance neutron peak and ARPES data in cuprates. *Phys. Rev. Lett.* **83**, 1652–1655 (1999).
- Zasadzinski, J. F. *et al.* Correlation of tunneling spectra in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ with the resonance spin excitation. *Phys. Rev. Lett.* **87**, 067005 (2001).
- Thomas, G. A. *et al.* $\text{Ba}_2\text{YCu}_3\text{O}_{7-x}$: Electrodynamics of crystals with high reflectivity. *Phys. Rev. Lett.* **61**, 1313–1316 (1988).
- Puchkov, A. V. *et al.* Evolution of the pseudogap state of high- T_c superconductors with doping. *Phys. Rev. Lett.* **77**, 3212–3215 (1996).
- Carbotte, J. P. *et al.* Coupling strength of charge carriers to spin fluctuations in high-temperature superconductors. *Nature* **401**, 354–356 (1999).
- Munzar, D., Bernhard, C. & Cardona, M. Possible relationship between the peak in magnetic susceptibility and the in-plane far infrared conductivity of YBCO. *Physica C* **317–318**, 547–549 (1999).
- Dai, P. *et al.* Resonance as a measure of pairing correlations in the high- T_c superconductor $\text{YBa}_2\text{Cu}_3\text{O}_{6.6}$. *Nature* **406**, 965–968 (2000).
- Demler, E. & Zhang, Z. C. Quantitative test of a microscopic mechanism of high-temperature superconductivity. *Nature* **396**, 733–735 (1998).
- Scalapino, D. J. The cuprate pairing mechanism. *Science* **284**, 1282–1283 (1999).
- Lanzara, A. *et al.* Evidence for ubiquitous strong electron-phonon coupling in high-temperature superconductors. *Nature* **412**, 510–514 (2001).
- Varma, C. M. *et al.* Phenomenology of the normal state of Cu-O high-temperature superconductors. *Phys. Rev. Lett.* **63**, 1996–1999 (1989).

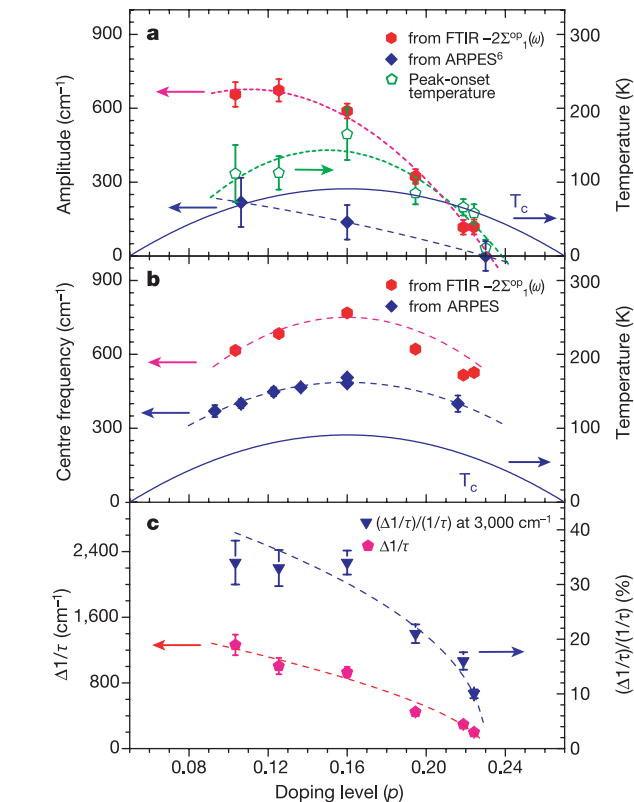


Figure 3 Doping-dependent properties of the optical resonance mode. See the Methods section for detailed descriptions. **a**, Amplitudes. **b**, Centre frequencies. The solid curve is $T_c(p)$ and the two dashed curves are parabolic fits, $5.6k_B T_c$ for ARPES and $8.0k_B T_c$ for FTIR. **c**, The quantity $\Delta 1/\tau$ (see Fig. 1a–d), the contribution of the resonance mode to the scattering rate, as a function of doping, and the same quantity normalized to the scattering rate at $3,000\text{ cm}^{-1}$. All the data point to the disappearance of the mode at a critical doping level of $p = 0.23$.

20. Kaminski, A. *et al.* Quasiparticles in the superconducting state of $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$. *Phys. Rev. Lett.* **84**, 1788–1791 (2000).
21. Millis, A. J. & Drew, H. D. Quasiparticles in high temperature superconductors: consistency of angle resolved photoemission and optical conductivity. Preprint at (<http://www.arXiv.org/cond-mat/0303018>) (2003).
22. Schachinger, E., Tu, J. J. & Carbotte, J. P. Angle-resolved photoemission spectroscopy and optical renormalizations: phonons or spin fluctuations. *Phys. Rev. B* **67**, 214508 (2003).
23. Dai, P. *et al.* The magnetic excitations spectrum and thermodynamics of high- T_c superconductors. *Science* **284**, 1344–1347 (1999).
24. Hwang, J. *et al.* Marginal Fermi liquid analysis of 300 K reflectance of $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$. Preprint at (<http://www.arXiv.org/cond-mat/0306250>) (2003).
25. Chakravarty, S. *et al.* Hidden order in the cuprates. *Phys. Rev. B* **63**, 094503 (2001).
26. Loram, J. W. *et al.* The condensation energy and pseudogap energy scale of $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ from electronic specific heat. *Physica C* **341–348**, 831–834 (2000).
27. Shibauchi, T. *et al.* Closing the pseudogap by Zeeman splitting in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ at high magnetic fields. *Phys. Rev. Lett.* **86**, 5763–5766 (2001).
28. Takenaka, K., Mizuhashi, K., Takagi, H. & Uchida, S. Interplane charge transport in $\text{YBa}_2\text{Cu}_3\text{O}_{7-y}$: Spin gap effect on in-plane and out-of-plane resistivity. *Phys. Rev. B* **50**, 6534–6537 (1994).
29. Homes, C. C. *et al.* Technique for measuring the reflectance of irregular, submillimeter-sized samples. *Appl. Opt.* **32**, 2976–2983 (1993).
30. Eisaki, H. *et al.* Effect of chemical inhomogeneity in the bismuth-based copper oxide superconductors. *Phys. Rev. B* (in the press); preprint at (<http://www.arXiv.org/cond-mat/0312429>) (2004).

Acknowledgements This work has been supported by the Canadian Natural Science and Engineering Research Council and the Canadian Institute of Advanced Research. We thank H. Eisaki and M. Greven for supplying us with several crystals. Their work at Stanford University was supported by the Department of Energy's Office of Basic Sciences, Division of Materials Science. The work at Brookhaven was supported in part by the Department of Energy. We thank D. N. Basov, J. P. Carbotte, G. M. Luke and M. R. Norman for discussions.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T.T. (timusk@mcmaster.ca).

Clarifying the glass-transition behaviour of water by comparison with hyperquenched inorganic glasses

Yuanzheng Yue¹ & C. Austen Angell²

¹Section of Chemistry, Department of Life Sciences, Aalborg University, DK-9000 Aalborg, Denmark

²Department of Chemistry and Biochemistry, Arizona State University, Tempe, Arizona 85287, USA

The formation of glasses is normal for substances that remain liquid over a wide temperature range (the 'good glassformers') and can be induced for most liquids if cooling is fast enough to bypass crystallization. During reheating but still below the melting point, good glassformers exhibit glass transitions as they abruptly transform into supercooled liquids, whereas other substances transform directly from the glassy to the crystalline state. Whether water exhibits a glass transition before crystallization has been much debated over five decades^{1,2–6}. For the last 20 years, the existence of a glass transition at 136 K (ref. 3) has been widely accepted^{2–4}, but the transition exhibits qualities difficult to reconcile with our current knowledge of glass transitions^{2,5,6}. Here we report detailed calorimetric characterizations of hyperquenched inorganic glasses that, when heated, do not crystallize before reaching their glass transition temperatures. We compare our results to the behaviour of glassy water and find that small endothermic effects, such as the one attributed to the glass transition of water, are only a 'shadow' of the real glass transition occurring at higher temperatures, thus substantiating

the conclusion⁶ that the glass transition of water cannot be probed directly.

Glassy water is the most abundant form of water in the Universe, found as thin films condensed on the interstellar dust particles that constitute the major component of comets⁷. This material, amorphous solid water (ASW), like hyperquenched glassy water (HQQW), is in a configurationally excited state (or state of high fictive temperature), relative to glasses formed by standard (20 K min^{−1}) cooling. To release the excitation enthalpy, the glass is usually annealed at temperatures where it does not crystallize. There is controversy about the state to which it is able to relax before crystallization occurs^{1–11}.

Until recently it was believed that ASW, HQGW and low-density amorphous water (LDA) all reach the internally equilibrated (or metastable liquid) state, so that it is a viscous liquid that is crystallizing. Observation of a weak endothermic effect at 136 K attributed to a glass transition³, penetration of millimetres-thick samples of LDA water (a state close in character to ASW) by a blunt probe at a temperature 14 K above the LDA glass transition temperature T_g of 129 K (ref. 8), and liquid-like diffusion at 150 K (ref. 9) support this view. Water, it was concluded, is a very fragile liquid, not only near its melting point¹⁰, but also near T_g (refs 9, 11).

The 136 K T_g is found by differential scanning calorimetry (DSC) using a heating rate of 10 K min^{−1}, but only after annealing the sample to avoid obscuring the endothermic signal by the release of the excitation enthalpy of the glass^{3,13,14}. (One important exception, where an endothermic effect is observed without prior sample annealing, is discussed below¹².) One difficulty with attributing the endotherm effect to the water T_g is that it is exceptionally weak: the jump in the heat capacity C_p of only 1.6 J mol^{−1} K^{−1} (ref. 14) is some 14 times smaller than the value expected at this temperature, according to extrapolations of data from unambiguous glass transitions in binary aqueous solutions that are good glassformers^{2,15}. This implies either that water undergoes a fragile-to-strong liquid transition in the experimentally inaccessible part of its phase diagram¹⁰, or that the 136 K endotherm is not caused by a glass transition at this temperature.

In this context, it is useful to consider that previous work failed to find evidence for a T_g endotherm^{5,16}; successful observation of the endotherm at 136 K required initial annealing of the glass for extended periods of 95 min (ref. 3) or 90 min (ref. 14), at about 6 K below the temperature at which the endothermic effect was subsequently identified. However, if the relaxation of the glassformer in question is strongly non-exponential, then annealing a quenched glass or even a standard glass far below its T_g is known to promote the appearance of a 'sub- T_g peak' during a subsequent DSC upscan^{17,18}. The shape of the observed endotherm at 136 K does in fact differ from the shape predicted for sub- T_g peaks using phenomenological models¹⁷; but modelling comparisons have not been made for hyperquenched glasses (HQGs) and the models are known to fail for systems far from equilibrium¹⁹.

From a recent comparison⁵ of the relaxation exotherm of unannealed HQGW with those of glassformers that do not crystallize it was concluded that the endotherm observed after annealing could not be due to a glass transition at 136 K, but could be due to relaxation of localized defects of the Bjerrum type known in ices. Here we advance an alternative, more generic, interpretation based on a careful calorimetric study of the behaviour of two HQGs and compared it with the endotherm behaviour of annealed HQGW.

Our glasses were both hyperquenched, by melt-spinning from ~2,200 K into ~8- μm fibres at 10⁶ K s^{−1} (see Methods), a rate comparable to that estimated for the hyperquenching of water³. One of our cases was chosen because of its extensive prior characterization^{20,21} and the second because it is simple in constitution ($\text{Al}_2\text{O}_3\text{-SiO}_2$) and, like water, crystallizes almost at T_g . Both are of very high T_g and hence are non-hygroscopic in character and are easily handled or stored.

To examine how glasses that are trapped in high-potential-energy states by hyperquenching recover equilibrium states that lie lower on the potential energy landscape²¹, we first characterize a 'standard glass' (a glass obtained by cooling at a rate of -0.33 K s^{-1} ; see also Fig. 1 caption). Then we compare, with this standard, the standard upscan of the HQG.

Figure 1 shows the thermal behaviour of the hyperquenched samples treated analogously to the HQGW samples, that is, upscanned at the standard rate after being annealed for 90 min at various temperatures well below T_g . The data show that if the annealing temperature (T_a) is sufficiently high, the upscan trace develops an endothermic component whose onset temperature appears to be fixed, but whose peak value depends on the temperature of annealing for a given annealing time (t_a).

In Fig. 2 we select one of these scans and compare it with the scan exhibiting the endotherm attributed in ref. 14 to the glass transition of water. The resemblance is striking, illustrating that it is possible to obtain a trace with the same form (and also the same small ΔC_p) as that attributed to the glass transition of water merely by appropriate annealing of the HQGW. Because the endotherm is determined by the relaxation characteristics of the material in the α -, or primary relaxation region, and is at least qualitatively predictable from the phenomenological models for the α -relaxation¹⁹, we rename it the 'shadow' glass transition with a transition temperature $T_{g,\text{shadow}}$ to distinguish it from any β -relaxation that would normally lie at a lower temperature with respect to the standard T_g .

For the crystallization-resistant mineral glass, the T_a can be placed arbitrarily close to the standard T_g , but, in the case of water, attempts to raise T_a much higher than 130 K for $t_a = 90 \text{ min}$ lead to crystallization. This feature explains why the shadow glass transition could be taken for the real transition.

In a recent study¹³, microdroplet water was deposited and vitrified at different substrate temperatures. Calorimetric characterization showed that whereas no endotherm appeared for water deposited at 130 K, an endotherm did appear for the sample

deposited at 140 K (see Fig. 3b, c). The 45-min deposition process itself establishes a temperature-time history amounting to an average 23-min anneal at the deposit temperature. The usual construction for the onset T_g (Fig. 1) now suggests that glassy water has a T_g higher than inferred before, namely 141 K. We demonstrate, in Fig. 3a, the parallel effect with hyperquenched glassy 35% Al_2O_3 -65% SiO_2 , which is only slightly more stable against crystallization than water.

Figure 3a shows the behaviour of this sample after long isochronal anneals at two different temperatures, 993 and 1,033 K, not far below the standard T_g . The sample annealed at 1,033 K exhibits a shadow peak that has risen to about a quarter the strength of the standard transition at overshoot. The crystallization of this glass can be seen to occur immediately after the overshoot peak of the standard scan. The transition temperature now attributed to the shadow glass transition has increased to 1,120 K, while annealing at 993 K yields a $T_{g,\text{shadow}}$ of only 1,063 K. This annealing-induced increase in T_g is 5%, compared to an increase of 4% in the case of water (141 K/136 K). A recent study¹⁸ asserted that the onset temperature of the sub- T_g peak is the same as the annealing temperature. This correlation does not hold for the $T_{g,\text{shadow}}$ value we obtain for the annealed HQGs (Fig. 3a), but it does seem to hold for the glassy water deposited and annealed at the higher temperature (Fig. 3b, c).

The behaviour of our two HQGs allows us to closely reproduce the behaviour of annealed HQGW, yet the endotherm of the glasses is only a 'shadow' of the real glass transition, which occurs some 15% higher in temperature. These findings strongly suggest that the water endotherm at 136 K is also only a 'shadow' of the real glass transition, substantiating earlier claims⁶ that the real T_g of water cannot be assigned except by some scaling procedure. These results parallel early studies on hyperquenched metallic glasses²² and have now also been obtained in studies on electrospray-quenched molecular glasses of dibutylphthalate and propylene glycol (L.-M. Wang and C.A.A., unpublished work).

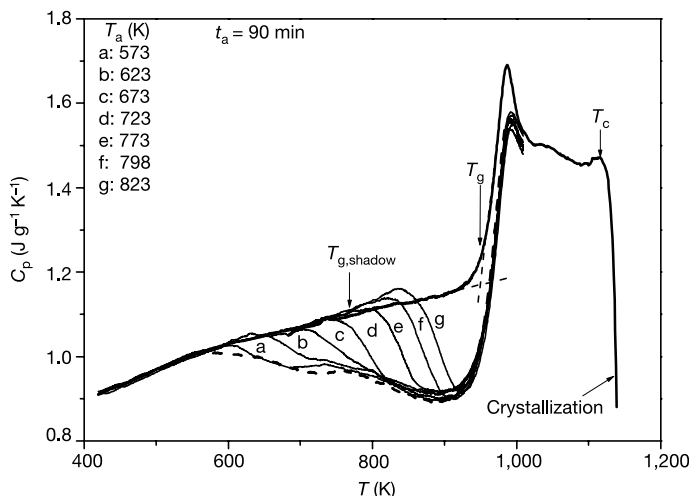


Figure 1 DSC upscans of hyperquenched and annealed HQGs. The thick lines show the comparison of the initial upscan of the hyperquenched glass at the standard rate (thick dashed line), with the upscan of the standard glass (thick solid line). The standard rate of 0.33 K s^{-1} was chosen because the 'onset-heating glass temperature' $T_{g,\text{onset}}$, defined here from a DSC standard upscan, is then the temperature at which the structural relaxation time is $\sim 100 \text{ s}$ (refs 2, 17). The upscan on the standard glass is continued until crystallization occurs at $T_c = 1,127 \text{ K}$. The standard upscan T_g is 944 K from the construction (dashed lines) shown. The quenched-in energy of the HQG is given by the area between the two curves, and is 64 J g^{-1} . The fictive temperature of the HQG,

determined as in refs 6 and 32, is 1,152 K. The thin solid lines (a–g) represent a series of standard upscans of aged samples. For each of lines a to g, a fresh sample of the HQG was annealed for 90 min at the temperature given, in order to observe the enthalpy relaxation of the remaining excess (frozen-in) enthalpy. Note, in particular, the evidence that the samples aged at temperatures above $\sim 723 \text{ K}$ reach lower enthalpies than the standard glass for fast-relaxing components of the structure, whereas the slow-relaxing components of the structure are still trapped in states of much higher enthalpy than that of the standard glass. The term $T_{g,\text{shadow}}$ is explained in the text.

Reproducibility of the endothermic effect¹⁴ has been one of the supporting arguments resulting in the general acceptance of 136 K as the T_g appropriate to water. We find that the 'sub- T_g peaks' seen with our inorganic glass over the temperature range 710–850 K (Fig. 1) can be reproduced any number of times by repeating the annealing process at the same temperature, provided that the crossover temperature (where endotherm becomes exotherm relative to the standard scan) is not exceeded on heating²⁰.

The endothermic effect is easily understood¹⁷. All glasses have a spectrum of relaxation times, as if composed of many microglasses of different T_g values. The structural relaxation of water (though not its dielectric relaxation) has a broad spectrum, as recently determined by optical Kerr-effect studies (R. Torre, personal communication). The shadow T_g is simply the collective T_g of that component of the HQG that has micro- T_g values low enough to be relaxed by the annealing.

Attributing the endothermic effect at 136 K to a sub- T_g peak (annealing pre-peak, or shadow glass transition) explains several perceived inconsistencies in the phase behaviour of glassy water. First, an activation energy (E_a) of 55.5 kJ mol⁻¹ has been inferred^{3,14,10} for the apparent glass transition endotherm, yet this value, when considered relative to T_g , is characteristic of a strong liquid^{2,10}, even though water in this temperature regime is thought to be a fragile liquid^{9,11}. Moreover, the E_a is close to the E_a values for relaxation processes in ice, as would be expected if the water is still a glass at 150 K. (Four independent studies of ASW crystallization obtained E_a values ranging from 55–84 kJ mol⁻¹, as reviewed in ref. 23.) Taken together, these considerations suggest that the E_a derived from the 136 K endotherm reflects the diffusion coefficient in the crystallizing phase²⁴. If water were a fragile liquid near T_g , this E_a should be >170 kJ mol⁻¹ (ref. 9).

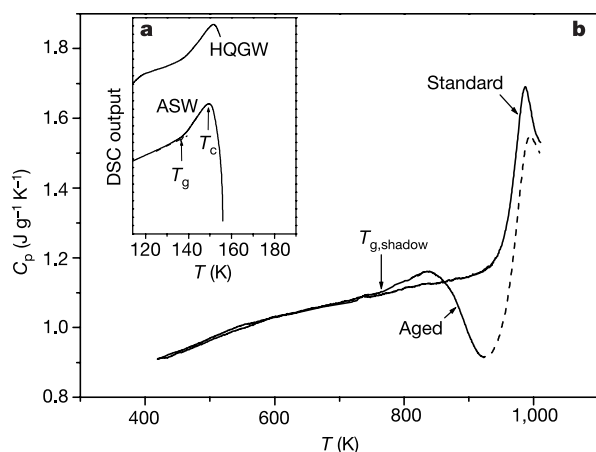


Figure 2 Comparison of DSC upscans of aged, hyperquenched mineral glass with the equivalent water DSC upscans for HQGW and ASW. The thermogram of the hyperquenched mineral glass sample, aged for 90 min at 823 K before upscanning (**b**), is compared with the upscans of HQGW and vapour-deposited ASW samples¹⁴ that had been aged for 90 min at 130 K before upscanning (**a**). The temperature scales are set so that each has 0 K at the same point offscale. The upscans for the water samples have provided the basis for assigning a T_g of 136 K to water. The comparison of **a** with **b** shows that the endotherm in **a** can arise from an annealing effect. The 'real' glass transition in the case of water has been eliminated from observation by crystallization. The dashed line in **b** is the part that we suggest has been cut from the water sample scan by crystallization, commencing at or slightly above the temperature marked T_c . The ratio of $T_{g,onset}$ to T_g of the mineral glass is 0.80. Applying the same ratio to $T_{g,shadow}$ of water we obtain a 'real but hidden' T_g of 169 K. However, this ratio will depend on system fragility, so the estimate of the hidden T_g for water is uncertain, and is better estimated by other means^{6,28}. (Data and identification on HQGW and ASW taken from ref. 10, re-plotted from ref. 14.)

As mentioned before, the extremely weak jump in C_p at T_g (see Fig. 2a and refs 3, 14) is 14-fold smaller than expected from extrapolations of ΔC_p data for seven different binary aqueous solutions^{2,15}. This apparent contradiction has been explained¹⁰ as being due to water changing from a fragile liquid to a strong liquid as temperature decreases during hyperquenching (or as the binary solutions change composition during dilution at low temperature). But it is not necessary to invoke a fragile-to-strong transition if the weak endotherm is simply an annealing pre-peak, rather than a real glass transition, given the ΔC_p values typically associated with 'shadow' and 'real' glass transitions (see Figs 1–3).

Attributing the long-controversial endothermic effect of glassy water to an annealing effect, not a T_g , also provides a simple explanation for puzzling changes²⁵ in the T_g of aqueous ethylene glycol (EG) and LiCl solutions, which first decreases and then suddenly increases by about 25 K with increase in EG or LiCl content. This sudden increase in T_g occurs on approach to the composition where the solutions become glassforming on normal cooling, but overshoots the published values owing to the 'normal' annealing effect¹⁷. (For the dilute compositions in ref. 25, an "optimal annealing temperature" had to be found to distinguish the EG-water ' T_g ' from 'sub- T_g ' shoulders observed on annealing a variety of glasses²⁵).

One argument for glassy water undergoing a glass transition to supercooled liquid is the observation that that LDA ($T_g = 129$ K, ref. 26) can be penetrated at 143 K by a blunt probe⁸ that penetrates *cis*-decalin ($T_g = 137.4$ K) at 140 K. The observation is puzzling, given that the dielectric relaxation (loss tangent) data of unannealed ASW, measured up to and through crystallization²⁷, clearly indicates that no relaxation occurs below 150 K relative to that for well-known glassformers like glycerol and propylene carbonate at their respective T_g values (see figure 4 of ref. 28). But the penetration temperature, 143 K, coincides with the temperature where crystallization is occurring for samples heated at the speed of the penetration measurement (compare with crystallization data in refs 26 and 29)—and mobility clearly increases during crystallization⁹.

To conclude, we note that the enthalpy relaxation plots and the

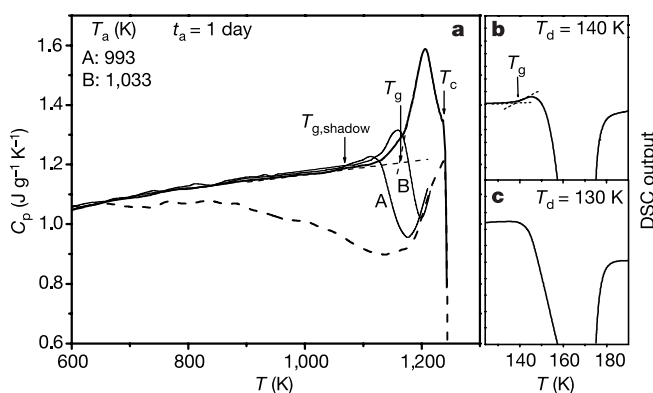


Figure 3 Effect of T_a on shadow glass transitions. **a**, DSC upscans of 35%Al₂O₃-65%SiO₂ standard glass (thick solid line) and HQG (dashed line) compared with two HQGs that had been annealed as in legend, showing shadow transitions (pre-peaks) at temperatures well above annealing temperatures. Shadow transitions increase in both transition strength and onset temperature with increase in T_a values. **b, c**, Upscans of vitreous water sample, deposited as 5- μ m droplets during 45 min, at temperatures given. The upscan of the deposit at $T_d = 130$ K (**c**) shows no endotherm before crystallization, while the deposit at $T_d = 140$ K (**d**) shows an endotherm with onset at 141 K by the usual construction (dashed lines). The shift from the original annealing-induced endotherm at 136 K (labelled T_g of water), is the analogue of the behaviour of the barely stable alumina-silica glass in curve B in **a**. (**b** and **c** are based on data presented in ref. 13.)

shadow glass transition of Fig. 1 give information on, respectively, frozen-in and annealed-out structural fluctuations. Therefore, the type of experiment shown in Fig. 1 may provide much information on water in its pre-vitrification supercooled liquid states. It is, after all, the large mean-square enthalpy fluctuations, responsible for the high heat capacity of supercooled water, that are being frozen-in near the fictive temperature at 200–230 K (ref. 30) during hyperquenching. Such hyperquench-and-anneal studies may also provide important information on water in different structural environments, such as around biomolecules, amphiphiles and the like. Indeed, the enthalpy relaxation and shadow T_g phenomenology may prove much more interesting to water studies than a mere glass transition. □

Methods

Sample preparation

The HQGs of this study were obtained by the cascade melt-spinning process. One sample is a multicomponent mineral (basalt) glass that has been well characterized elsewhere²⁰: composition (mol%): 47.5 SiO₂, 9.1 Al₂O₃, 6.6 FeO, 15.9 CaO, 16.9 MgO, 2.0 Na₂O, 0.7 K₂O. The second is from a binary aluminosilicate melt with the composition (mol%): 65 SiO₂, 35 Al₂O₃. Before the spinning, the melts were homogenized at temperatures of about 1,800–1,850 K (for the mineral system) and about 2,173–2,223 K (for the binary system) for half an hour. The fibres were spun using the wheel centrifuge process, which has also been called the cascade or mechanical spinning process³¹. The glass melt is conveyed down a trough and onto the outer rim of a spreader wheel. Although some material is spun off, most of the melt is transferred to either of two larger adjacent wheels that are spinning at 6,000 r.p.m. in the opposite direction. Fibres form when droplets of diameter >65 µm are thrown from the wheels by centripetal force. To remove the remaining glass droplets from the fibre ends, and to ensure uniform fictive temperature and properties, the fibres were separated using a 65-µm sieve. Of the fibre volume, 84 vol.% is of diameter <12.6 µm, 50 vol.% is of $d < 7.7$ µm, and only 16 vol.% have $d < 4.1$ µm. The average hyperquenching rate was estimated to be 2×10^6 K s⁻¹, by calculation from the relation between cooling rate Q and shear viscosity η at the fibre fictive temperature: $\log(1/Q) = \log \eta - 11.35$ (Y.-Z.Y., R. von der Ohe, & S. L. Jensen, unpublished work).

Ageing and calorimetric experiments

Before the DSC measurements, the fibre samples were annealed in an annealing furnace at selected temperatures T_a for $t_a = 90$ min (see Figs 1 and 2), whereas the binary samples were annealed at selected temperatures for 24 h (see Fig. 3). The annealing was performed in air. The apparent C_p of each fibre sample was measured using a Netzsch STA449C DSC. The fibres were placed in a platinum crucible situated on a sample holder of the DSC at room temperature. The samples were held for 5 min at an initial temperature of 333 K, and then heated at a standard rate of 0.333 K s⁻¹ to the temperature 1,013 K (for the mineral samples) and 1,236 K (for the binary samples), and then cooled back to 573 K at a rate of 0.333 K min⁻¹ to 573 K, thus forming the standard glass. After natural cooling to room temperature, the second upscan was performed using the same procedure as for the first. To determine the C_p of the fibres, both the baseline (blank) and the reference sample (sapphire) were measured. To confirm reproducibility, the measurements for some samples were repeated to check for drift in the baseline.

Received 3 July; accepted 12 December 2003; doi:10.1038/nature02295.

1. Pryde, J. A. & Jones, G. O. Properties of vitreous water. *Nature* **170**, 635–639 (1952).
2. Angell, C. A. Liquid fragility and the glass transition in water and aqueous solutions. *Chem. Rev.* **102**, 2627–2649 (2002).
3. Johari, G. P., Hallbrucker, A. & Mayer, E. The glass transition of hyperquenched glassy water. *Nature* **330**, 552–553 (1987).
4. Johari, G. P. Does water need a new T_g ? *J. Chem. Phys.* **116**, 8067–8073 (2002).
5. MacFarlane, D. R. & Angell, C. A. Nonexistent glass transition for amorphous solid water. *J. Phys. Chem.* **88**, 759–762 (1984).
6. Velikov, V., Borick, S. & Angell, C. A. The glass transition of water, based on hyperquenching experiments. *Science* **294**, 2335–2338 (2001).
7. Jenniskens, P., Banham, S. F., Blake, D. F. & McCoustra, M. R. S. Liquid water in the domain of cubic crystalline ice I-c. *J. Chem. Phys.* **107**, 1232–1241 (1997).
8. Johari, G. P. Liquid state of low-density pressure-amorphized ice above its T_g . *J. Phys. Chem. B* **102**, 4711–4714 (1998).
9. Smith, R. S. & Kay, B. D. The existence of supercooled liquid water at 150 K. *Nature* **398**, 788–791 (1999).
10. Ito, K., Moynihan, C. T. & Angell, C. A. Thermodynamic determination of fragility in liquids and a fragile-to-strong liquid transition in water. *Nature* **398**, 492–495 (1999).
11. Kivelson, D. & Tarjus, G. H₂O below 277 K: A novel picture. *J. Phys. Chem. B* **105**, 6620–6627 (2001).
12. Hallbrucker, A. & Mayer, E. Calorimetric study of the vitrified liquid water to cubic ice phase transition. *J. Phys. Chem.* **91**, 503–505 (1987).
13. Kohl, I., Mayer, E. & Hallbrucker, A. The glassy water-cubic ice system: a comparative study by X-ray diffraction and differential scanning calorimetry. *Phys. Chem. Chem. Phys.* **2**, 1579–1586 (2000).
14. Hallbrucker, A., Mayer, E. & Johari, G. P. Glass-liquid transition and the enthalpy of devitrification of annealed vapor-deposited amorphous solid water. A comparison with hyperquenched glassy water. *J. Phys. Chem.* **93**, 4986–4990 (1989).
15. Angell, C. A. & Tucker, J. C. Heat capacity changes in glass-forming aqueous solutions, and the glass transition in vitreous water. *J. Phys. Chem.* **84**, 268–272 (1980).

16. Ghormley, J. A. Enthalpy changes and heat-capacity changes in transformations from high-surface-area amorphous ice to stable hexagonal ice. *J. Chem. Phys.* **48**, 503–511 (1968).
17. Hodge, I. M. Enthalpy recovery in amorphous materials. *J. Non-Cryst. Solids* **169**, 211–266 (1994).
18. Johari, G. P. Water's endotherm, the sub- T_g peak of glasses and T_g of water. *J. Chem. Phys.* **119**, 2935–2937 (2003).
19. Huang, J. & Gupta, P. Enthalpy relaxation in thin glass fibers. *J. Non-Cryst. Solids* **151**, 175–181 (1992).
20. Yue, Y.-Z., Jensen, S. L. & Christiansen, J. de C. Physical aging in a hyperquenched glass. *Appl. Phys. Lett.* **81**, 2983–2985 (2002).
21. Angell, C. A. *et al.* Potential energy, relaxation, vibrational dynamics and the boson peak, of hyperquenched glasses. *J. Phys. Condensed Matter* **15**, S1051–S1068 (2003).
22. Inoue, A., Masumoto, T. & Chen, H. S. Enthalpy relaxation behaviour of metal-metal (Zr-Cu) amorphous alloys upon annealing. *J. Mater. Sci.* **20**, 4057–4068 (1985).
23. Angell, C. A. in *Water Science for Food, Health, Agriculture and Environment* (eds Berk, Z., Leslie, R. B., Lilford, P. J. & Mizrahi, S.) 1–30 (Technomic, Lancaster, 2001).
24. Ngai, K. L., Magill, J. H. & Plazek, D. J. Flow, diffusion and crystallization of supercooled liquids: Revisited. *J. Chem. Phys.* **112**, 1887–1892 (2000).
25. Hofer, K., Hallbrucker, A., Mayer, E. & Johari, G. P. Vitrified dilute aqueous-solutions. 3. Plasticization of waters H-bonded network and the glass-transition temperatures minimum. *J. Phys. Chem.* **93**, 4674–4677 (1989).
26. Johari, G. P., Hallbrucker, A. & Mayer, E. Two calorimetrically distinct states of liquid water below 150 kelvin. *Science* **273**, 90–92 (1996).
27. Johari, G. P., Hallbrucker, A. & Mayer, E. The dielectric behavior of vapor-deposited amorphous solid water and of its crystalline forms. *J. Chem. Phys.* **95**, 6849–6855 (1991).
28. Angell, C. A. Amorphous water. *Annu. Rev. Phys. Chem.* **55**, 559–583 (2004).
29. Handa, Y. P. & Klug, D. D. Heat capacity and glass transition of amorphous ice. *J. Chem. Phys.* **92**, 3323–3325 (1988).
30. Fleissner, G., Hallbrucker, A. & Mayer, E. Increasing contact-ion pairing as a supercooled water anomaly. Estimation of the fictive temperature of hyperquenched glassy water. *J. Phys. Chem. B* **102**, 6239–6247 (1998).
31. Axten, C. W. *et al.* *Man-made Vitreous Fibers: Nomenclature, Chemical and Physical Properties* (ed. Easters, W.) 17 (TIMA, Inc., Stamford, CT, 1991–1993).
32. Yue, Y.-Z., Christiansen, J. de C. & Jensen, S. L. Determination of the fictive temperature for a hyperquenched glass. *Chem. Phys. Lett.* **357**, 20–24 (2002).

Acknowledgements We are grateful for the support of Rockwool International, Denmark, Department of Production, Aalborg University, Denmark (to Y.-Z.Y.) and for a Solid State Chemistry grant from the National Science Foundation (C.A.A.). We thank S.L. Jensen for help with samples.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to C.A.A. (caa@asu.edu).

High-latitude influence on the eastern equatorial Pacific climate in the early Pleistocene epoch

Zhonghui Liu & Timothy D. Herbert

Department of Geological Sciences, Brown University, Providence, Rhode Island 02912, USA

Many records of tropical sea surface temperature and marine productivity exhibit cycles of 23 kyr (orbital precession) and 100 kyr during the past 0.5 Myr (refs 1–5), whereas high-latitude sea surface temperature records display much more pronounced obliquity cycles at a period of about 41 kyr (ref. 6). Little is known, however, about tropical climate variability before the mid-Pleistocene transition about 900 kyr ago, which marks the change from a climate dominated by 41-kyr cycles⁷ (when ice-age cycles and high-latitude sea surface temperature variations were dictated by changes in the Earth's obliquity^{8,9}) to the more recent 100-kyr cycles of ice ages. Here we analyse alkenones from marine sediments in the eastern equatorial Pacific Ocean to reconstruct sea surface temperatures and marine productivity over the past 1.8 Myr. We find that both records are dominated by the 41-kyr

obliquity cycles between 1.8 and 1.2 Myr ago, with a relatively small contribution from orbital precession, and that early Pleistocene sea surface temperatures varied in the opposite sense to local annual insolation in the eastern equatorial Pacific Ocean. We conclude that during the early Pleistocene epoch, climate variability at our study site must have been determined by high-latitude processes that were driven by orbital obliquity forcing.

Since Milankovitch suggested that the ice ages were controlled by summer insolation variations at northern middle and high latitudes¹⁰, studies of marine palaeoclimates have shown that this theory accounts for much of the climate variance of the late Pleistocene time¹¹, with the as-yet unexplained addition of substantial spectral power at 100 kyr (ref. 12). High-latitude seasonal insolation is controlled to a very large degree by ~21-kyr precession cycles, and the low-latitude seasonal extremes are entirely dominated by precession¹³. Over the annual cycle, however, precessional insolation anomalies sum to zero at all latitudes. In contrast, obliquity variations (dominant period at 41 kyr) change the annual insolation received as a function of latitude. A high-obliquity configuration increases insolation at latitudes above 45° in both hemispheres and decreases the solar energy received in the tropics and subtropics. The effects of obliquity are quite small relative to mean annual insolation: for example, about 1.5% ($\pm 3.3 \text{ W m}^{-2} \text{ yr}^{-1}$) at 65° N, and only about 0.4% ($\pm 1.7 \text{ W m}^{-2} \text{ yr}^{-1}$) at 3° S, the latitude of our study site. Significant amplifiers in the atmosphere

and ocean (moisture and heat transport, ice-albedo, sea-ice cover) would therefore be needed to transform the obliquity forcing into large-scale climate change.

Most of our knowledge of the evolution of Pleistocene palaeoclimates comes from oxygen isotopic (ice volume) data. Before the mid-Pleistocene transition, ice volume oscillated on a 41-kyr timescale dictated by obliquity insolation cycles⁸. The amplitude of the 41-kyr component in benthic $\delta^{18}\text{O}$ varies little over the past 1.8 Myr, while the 23-kyr, and particularly the 100-kyr, components of ice volume variance have grown significantly in the late Pleistocene. In contrast to the ice volume signal, which represents a global integration, late Pleistocene sea surface temperatures (SSTs) in the North Atlantic display a latitudinal pattern with relatively more 41-kyr power at high latitudes, and more 23-kyr power at low latitudes⁶. Thus, 41-kyr rhythms are believed to be a 'polar' signal. Aspects of the tropical surface ocean, including indices of surface water biological production and monsoonally related changes in wind intensity, in most cases, show pronounced 23-kyr power in the late Pleistocene^{1–5}, despite the strong 41-kyr power also present in some records^{14,15}. These aspects of tropical surface oceanography may reflect changes in the local wind field, which controls the intensity of upwelling, and the more remote forcing of the tropical thermocline. In light of the data presented below, it is important to note that late Pleistocene time series representative of tropical wind fields do not overwhelmingly resemble SST (see, for example, ref. 15 versus ref. 16).

Very little is known about how the surface ocean, especially in the tropics, operated during the time interval dominated by 41-kyr glacial cycles^{9,17}. We characterized SST and palaeoproductivity variance during the heart of the '41-kyr world' at Ocean Drilling Program (ODP) Site 846 (3° 5' S, 90° 49' W), using the alkenone

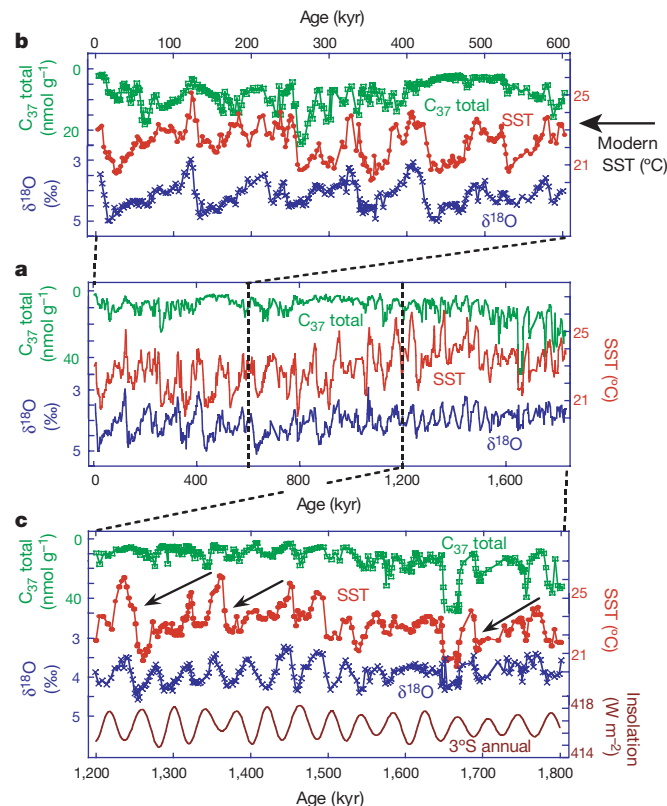


Figure 1 Time series of alkenone abundances (C_{37} total), SST, ice volume¹⁹ (benthic $\delta^{18}\text{O}$) and local annual insolation at ODP Site 846. **a**, The entire Pleistocene (0–1.8 Myr ago) time series. **b**, Detailed view of late Pleistocene (0–0.6 Myr ago) time series. **c**, Detailed view of early Pleistocene (1.2–1.8 Myr ago) time series. Alkenone abundance is plotted inversely (C_{37} total increases with cold SST). Arrow lines in **c** indicate saw-tooth cycles in the SST record. Note that annual local insolation at 3° S in the early Pleistocene, obtained from Analyseries software²⁷, is anti-phased to SST.

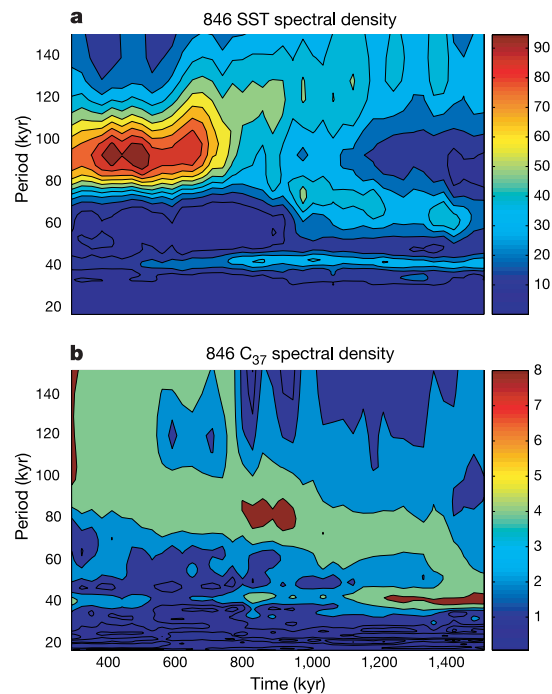


Figure 2 Evolution of spectral densities in the entire Pleistocene (0–1.8 Myr ago). **a**, SST. **b**, Alkenone abundances (log-transformed, see Methods). The length of time window for spectral analysis is 600 kyr and the moving step is 30 kyr, thus the first and last data points represent the intervals 0–0.6 Myr ago and 1.23–1.83 Myr ago, respectively. Spectral analyses are done using the Blackman–Tukey method, lagged by 150 data points in a time window of 300 data values. Units of spectral density: **a**, $^{\circ}\text{C}^2 \text{ kyr}^{-1}$; **b**, $(\text{nmol g}^{-1})^2 \text{ kyr}^{-1}$.

organic proxy (see Methods). Our study site is located in the eastern equatorial Pacific (EEP) cold tongue¹⁸. Here, the Peru Current merges with the Southern Equatorial Current (Supplementary Fig. S1). Relatively high nutrient levels in the euphotic zone, high biological production, and cold SSTs result from the shoaling of the Equatorial Undercurrent and advection of water from the eastern boundary current along the Peru–Chile margin. The eastern boundary currents and the Southern Equatorial Current are strongly influenced by the changes in the atmospheric circulation of the Southern Hemisphere (trade winds) on seasonal and interannual timescales.

The benthic $\delta^{18}\text{O}$ record at Site 846¹⁹ provides an orbitally tuned timescale, and allows us to compare SST and palaeoproductivity records to the global ice volume signal for the entire Pleistocene (see Methods). For the past 1.25 Myr, our SST record is broadly similar to a previous low-resolution alkenone-SST record produced at the same site²⁰. The alkenone-derived SST followed benthic $\delta^{18}\text{O}$ very closely, with warmer temperatures during periods of lower ice volume (Fig. 1a), and recorded a temperature range (up to 4.5 °C) in the early Pleistocene similar to that of the last 500 kyr (Fig. 1b, c). Glacial–interglacial temperature differences in the early Pleistocene varied from about 1 °C to 4.5 °C. Although temperatures in the interval 1.2–1.8 Myr ago were about 1 °C higher than the average late Pleistocene SST, the temperature did not change monotonically at Site 846.

In addition to providing estimates of past SST, alkenone concentrations (C_{37} total) provide a proxy for past trade wind strength

and thermocline depth. Where the records overlap, alkenone abundances at Site 846 strongly resemble a nearby record of past thermocline depth based on the abundance of the nanofossil *Florisphaera profunda*¹⁵ (C_{37} total is high when *F. profunda* is low), and they also correlate positively with organic carbon and carbonate accumulation at Site 846²⁰. Alkenone abundances in the interval 1.2–1.8 Myr ago were higher than those in the more recent sediments. Early Pleistocene alkenone production was consistently high in glacial periods (2–15 times more than that of interglacial periods). The association between high C_{37} total and cold SST declined from the early Pleistocene to the late Pleistocene ($r^2 = 0.46$ and $r^2 = 0.04$, respectively). Within the early Pleistocene, temperatures were coolest and alkenone production was highest in the interval 1.5–1.8 Myr ago. A similar cold phase appears to have been detected in the planktonic $\delta^{18}\text{O}$ record at ODP Site 677, slightly north of the Equator in the EEP⁸.

Evolutionary spectral analyses of SST and productivity proxies show that 41-kyr power predominated in the early Pleistocene, and 100-kyr power in the late Pleistocene (Fig. 2). The amplitude of the early Pleistocene obliquity cycle in SST and productivity was nearly twice that of the past 600 kyr, but the tilt cycle outweighed precession even in the late Pleistocene (Supplementary Fig. S2). In the early Pleistocene, lower-frequency spectral components appeared at integer multiples (about 80 and 120 kyr) of the obliquity period; these can be seen as saw-tooth groupings of two or three obliquity cycles in the 1.2–1.8 Myr SST time series (Fig. 1c). Tropical SST and productivity were highly coherent with ice volume in the obliquity band over the entire Pleistocene. Both SST and productivity led ice volume by 2–3(± 1) kyr in the 1.2–1.8 Myr interval (Fig. 3). However, in the interval 0–0.6 Myr ago, SST and productivity led ice volume in the obliquity band by 5 ± 1 kyr and by 3 ± 2 kyr, respectively (Supplementary Fig. S3). A similar phase relationship in the mid- and late Pleistocene is also identified in a previous study²⁰. In the precession band, sea surface signals in the EEP were less coherent with ice volume than in the obliquity band. Only early Pleistocene productivity and late Pleistocene SST in the precession band were coherent with ice volume at the 80%, but not the 95%, level, and spectral power at precessional (~23 kyr) frequencies was small throughout the Pleistocene SST and productivity records (Fig. 2).

During the glacial phase of the 41-kyr cycles, the simplest mechanism that could produce both cold equatorial SSTs and high productivity involves a shoaling of the thermocline and enhanced trade winds. A cold SST does not by itself imply increased upwelling, because cold temperatures could be advected from elsewhere. However, the strong correlation between cold SST at Site 846 and increases in the alkenone content of the sediment support upwelling as at least one important process that controlled SST in the study region. Today, the large tongue of advected cold water in the EEP is almost exclusively the result of Southern Hemisphere wind patterns²¹. The early Pleistocene 41-kyr cycles, involving both SST and productivity, suggest that equatorial trade winds, as well as the larger-scale heat loop, varied in concert with high-latitude processes. In the late Pleistocene, the weaker coupling between cold SSTs and high productivity indicates that upwelling was a less important factor controlling SST during this interval.

The predominant 41-kyr climatic cycles in the early Pleistocene could not arise from a direct tropical response to orbital forcing, seasonally or annually, because (1) local seasonal insolation was completely dominated by precession cycles, and (2) local annual insolation peaked when SST was cool (Fig. 1c). The local seasonal insolation might however account for the precessional power in the SST and productivity records by changing the slope of the Pacific east–west thermocline¹³. The smaller coherence between EEP surface proxies and ice volume in the precession as compared with the obliquity band might indicate that the precessional response in the

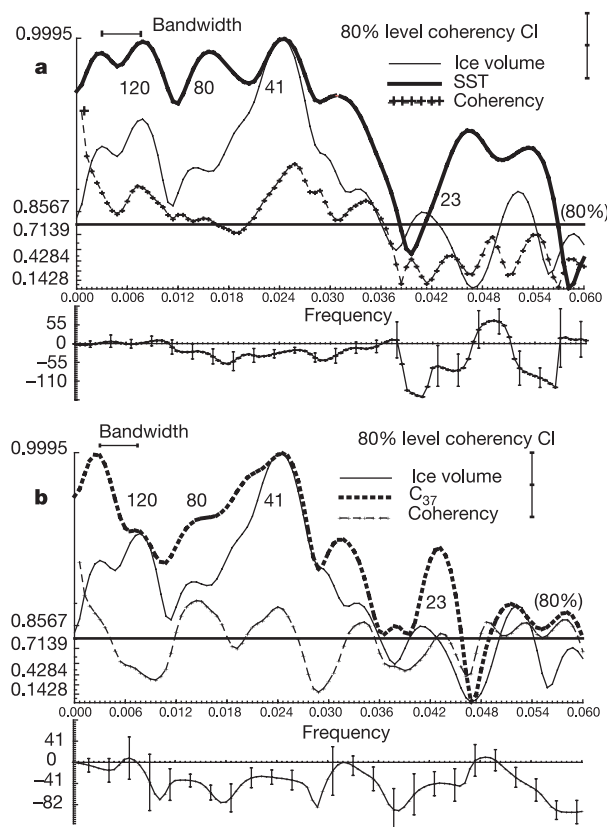


Figure 3 Coherencies and phase relationships in the early Pleistocene (1.2–1.8 Myr ago). **a**, SST with respect to ice volume. **b**, Alkenone abundance (log-transformed) with respect to ice volume. Spectral densities of SST and alkenone abundance are normalized and plotted on log scales. Orbital periods (41 and 23 kyr) and their subharmonics (~80 and 120 kyr) are indicated. Negative values in phase plots indicate that SST or alkenone abundance leads ice volume. CI, confidence interval.

EEP was caused by local precessional forcing, and thus less coupled to the high-latitude ice volume (we note that the smaller coherency could be also due to the very small precessional power with the presence of random noise in the time series). However, the majority of spectral power in our early Pleistocene SST and productivity records is located at the obliquity period and its subharmonics. Therefore, a remote connection between high latitudes and the EEP must account for the early Pleistocene surface signals detected. Extra-tropical conditions, such as the distribution and orography of ice sheets, the land–sea distribution, and the Equator to pole temperature gradient, may affect the tropical atmospheric circulation pattern²². One might expect that ice-albedo feedbacks to global climate would have been less significant in the early as compared with the late Pleistocene, owing to the smaller size of Northern Hemisphere ice sheets, and, therefore, that tropical and high-latitude conditions were not tightly locked. We observe instead that tropical SST fluctuated closely with ice volume, and in phase with insolation at high, not low, latitudes. The obliquity response in the EEP in the early Pleistocene, which was coherent with and slightly led ice volume, might suggest that Northern Hemisphere ice sheets did not provide the initial forcing of tropical surface water variability, but they could have provided a strong ice-albedo feedback. Two possible remote connections could closely couple early Pleistocene tropical SST and wind patterns to the obliquity forcing at high latitudes via atmospheric circulation⁷ or thermocline circulation²³.

The pattern of SST and palaeoproductivity changes derived from ODP Site 846 suggest that climatic evolution during the Pleistocene was considerably more complex than suggested by the constancy of the 41-kyr component of benthic $\delta^{18}\text{O}$ throughout the Pleistocene⁸. The coupling between low- and high-latitude SST may have been tighter under the smaller ice sheet conditions of the early Pleistocene, when the climate was strongly tuned to obliquity forcing. However, eastern equatorial Pacific SSTs and trade winds became more sensitive to the combination of precession/eccentricity over the past 500 kyr (refs 1–5), although the obliquity signal persisted in a weakened form. Our results will need to be corroborated by data from other tropical locations, but they characterize at a minimum the variance of the largest open-ocean upwelling system in the world during the 41-kyr period of Earth's history. □

Methods

Alkenone palaeothermometry

Alkenones are produced by certain species of haptophyte (coccolithophorid) algae that live in the near-surface and faithfully record growth temperatures^{24,25}, which we here equate with mean annual SST. A global study of recent sediments found that the highest correlation between alkenone unsaturation index (U_{37}^k) and near-surface temperatures came from comparing the U_{37}^k index to annual mean temperature at the sea surface²⁶. This data set includes samples from regions that are today dominated by *Gephyrocapsa oceanica*, the most likely alkenone-producing haptophyte species prior to the evolutionary appearance of *Emiliania huxleyi* during oxygen isotope stage 8 (refs 25, 26). Because the U_{37}^k values of modern sediments dominated by *Gephyrocapsa* production fall on the global regression to mean annual SST, we assume that the modern calibration²⁶ applies to all values of the unsaturation index determined for ODP Site 846. Alkenone extraction and analysis by gas chromatography in our laboratory followed procedures detailed elsewhere²⁵. Our analyses also quantify alkenone abundances in the sediment, an indicator of coccolithophorid productivity^{3,4}.

Age model

The Site 846 age model¹⁹ is tuned to the SPECMAP timescale for the interval 0–0.6 Myr ago and to the ODP Site 677 timescale⁶ for the interval 0.6–1.8 Myr ago. In fact, there is little difference in the 0–0.6 Myr interval between the ages of isotopic stages in the SPECMAP and the 677 timescales⁸. Thus the potential effects of combining two different timescales for the 846 age model are minimal. The Site 677 timescale is indirectly tuned to precession, whereas the SPECMAP timescale is tuned to all the orbital parameters in order to obtain the best fit. Thus, if there was any difference caused by combining two age models for two time intervals, the spectral bias would be in the sense of increasing obliquity power in the 0–0.6 Myr interval relative to the early Pleistocene. We observed instead that the strongest obliquity signal occurred in the SST and productivity records at 1.2–1.8 Myr ago. In this study, we have found that the tropical SST in the early Pleistocene is anti-phased to the local annual insolation (Fig. 1c). Although tuning often assumes a certain lag between an insolation target and a benthic $\delta^{18}\text{O}$ record, the maximum realistic lag for the

precessionally tuned age model used here is insufficient to explain this anti-phased relationship. Furthermore, the phase relationships among $\delta^{18}\text{O}$, SST and alkenone abundance, obtained from the same site, are independent of tuning.

Spectral analysis

In Fig. 2b, we performed evolutionary spectral analysis after taking the logarithm of alkenone abundances because the distribution of the entire data set is lognormal, rather than gaussian. Spectra of the log-transformed alkenone abundances in the early Pleistocene reveal the obliquity power centred precisely at 41 kyr. However, spectra of linear alkenone abundances show the obliquity peak shifted to a slightly lower frequency, although alkenone abundances were tightly coupled to SST and ice volume, whose obliquity power is precisely centred at 41 kyr.

Received 17 July 2003; accepted 12 January 2004; doi:10.1038/nature02338.

- Molfini, B. & McIntyre, A. Precessional forcing of nutrient dynamics in the Equatorial Atlantic. *Science* **249**, 766–769 (1990).
- Schneider, R. R., et al. in *The South Atlantic: Present and Past Circulation* (eds Wefer, G., Berger, W. H., Siedler, G. & Webb, D. J.) 527–551 (Springer, Berlin, 1996).
- Rostek, E., Bard, E., Beaufort, L., Sonzogni, C. & Ganssen, G. Sea surface temperature and productivity records for the past 240 kyr in the Arabian Sea. *Deep Sea Res.* **44**, 1461–1480 (1997).
- Budziak, D. S. et al. Late Quaternary insolation forcing on total organic carbon and C_{37} alkenone variations in the Arabian Sea. *Paleoceanography* **15**, 307–321 (2000).
- Perks, H., Charles, C. D. & Keeling, R. F. Precessionally forced productivity variations across the equatorial Pacific. *Paleoceanography* **17**, 63–69 (2002).
- Ruddiman, W. F. & McIntyre, A. Ice-age thermal response and climatic role of the surface North Atlantic ocean, 40 to 63 N. *Bull. Geol. Soc. Am.* **95**, 381–396 (1984).
- Raymo, M. E. & Nisancioglu, K. The 41 kyr world: Milankovitch's other unsolved mystery. *Paleoceanography* **18**, 10.1029/2002PA000791 (2003).
- Shackleton, N. J., Berger, A. & Peltier, W. R. An alternative astronomical calibration of the lower Pleistocene timescale based on ODP site 677. *Trans. R. Soc. Edinb. Earth Sci.* **81**, 251–261 (1990).
- Ruddiman, W. F., Raymo, M. E., Martinson, D. G., Clement, B. M. & Backman, J. Pleistocene evolution: northern hemisphere ice sheets and North Atlantic Ocean. *Paleoceanography* **4**, 353–412 (1989).
- Milankovitch, M. K. Kanon der erdbestrahlung und seine anwendung auf daseiszeitenproblem. *Serb. Acad. Beogr. Spec. Publ.* **132** (1941).
- Hays, J. D., Imbrie, J. & Shackleton, N. J. Variations in the Earth's orbit: Pacemaker of the ice ages. *Science* **194**, 1121–1132 (1976).
- Imbrie, J. et al. On the structure and origin of major glaciation cycles 2: The 100,000 year cycle. *Paleoceanography* **8**, 699–735 (1993).
- Berger, A. L. Long-term variations of daily insolation and quaternary climatic change. *J. Atmos. Sci.* **35**, 2362–2367 (1978).
- Clemens, S., Prell, W., Murray, D., Shimmield, G. & Weedon, G. Forcing mechanisms of the Indian ocean monsoon. *Nature* **353**, 720–725 (1991).
- Beaufort, L., de Garidel-Thoron, T., Mix, A. C. & Pisias, N. G. ENSO-like forcing on oceanic primary production during the late Pleistocene. *Science* **293**, 2440–2444 (2001).
- Lea, D. W., Pak, D. K. & Spero, H. J. Climate impact of late Quaternary equatorial Pacific sea surface temperature variations. *Science* **289**, 1719–1724 (2000).
- Bloemendal, J. & deMenocal, P. Evidence for a change in the periodicity of tropical climate cycles at 2.4 Myr from whole-core magnetic susceptibility measurements. *Nature* **342**, 897–900 (1989).
- Levitus, S. & Boyer, T. P. *World Ocean Atlas 1994 Vol. 4, Temperature* (NOAA Atlas NESDIS Vol. 4, US Department of Commerce, Washington DC, 1994).
- Mix, A. C., Le, J. & Shackleton, N. J. Benthic foraminiferal stable isotope stratigraphy of Site 846: 0–1.8 Ma. *Sci. Res. ODP* **138**, 839–856 (1995).
- Emeis, K.-C., Dooze, H., Mix, A. & Schulz-Bull, D. Alkenone sea-surface temperatures and carbon burial at Site 846 (eastern equatorial Pacific Ocean): the last 1.3 m.y. *Sci. Res. ODP* **138**, 605–614 (1995).
- Lyle, M. W., Pahl, F. G. & Sparrow, M. A. Upwelling and productivity changes inferred from a temperature record in the central equatorial Pacific. *Nature* **355**, 812–815 (1992).
- Bush, A. B. G. & Philander, S. G. H. The role of ocean-atmosphere interactions in tropical cooling during the last glacial maximum. *Science* **279**, 1341–1344 (1998).
- Philander, S. G. H. & Fedorov, A. V. Role of tropics in changing the response to Milankovitch forcing some three million years ago. *Paleoceanography* **18**, 10.1029/2002PA000837 (2003).
- Pahl, F. G., Muehlhausen, L. A. & Zahnle, D. L. Further evaluation of long-chain alkenones as indicators of paleoceanographic conditions. *Geochim. Cosmochim. Acta* **52**, 2303–2310 (1988).
- Herbert, T. D. et al. Depth and seasonality of alkenone production along the California margin inferred from a core-top transect. *Paleoceanography* **13**, 263–271 (1998).
- Muller, P. J., Krist, G., Ruhland, G., Von Storch, I. & Rosell-Mele, A. Calibration of the alkenone paleotemperature index U_{37}^k based on core-tops from the eastern South Atlantic and the global ocean (60N–60S). *Geochim. Cosmochim. Acta* **62**, 1757–1772 (1998).
- Paillard, D., Labeyrie, L. & Yiou, P. Macintosh program performs time-series analysis. *Eos* **77**, 379 (1996).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank the curators of the Ocean Drilling Program (Texas A&M University) for providing samples. The original draft was improved by comments from S. Clemens, K. Lawrence and L. Lisiecki. This work was supported by the NSF (T.D.H.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to Z.L. (Zhonghui_Liu@brown.edu).

Aftershocks driven by a high-pressure CO₂ source at depth

Stephen A. Miller¹, Cristiano Collettini², Lauro Chiaraluce³, Massimo Cocco³, Massimiliano Barchi² & Boris J. P. Kaus⁴

¹Institute of Geophysics, Swiss Federal Institute of Technology (ETH), 8093 Zürich, Switzerland

²Università degli Studi di Perugia, Perugia, 06100 Italy

³Istituto Nazionale di Geofisica e Vulcanologia, Rome, 00143 Italy

⁴Geology Institute, Swiss Federal Institute of Technology (ETH), 8092 Zürich, Switzerland

In northern Italy in 1997, two earthquakes of magnitudes 5.7 and 6 (separated by nine hours) marked the beginning of a sequence that lasted more than 30 days, with thousands of aftershocks including four additional events with magnitudes between 5 and 6. This normal-faulting sequence is not well explained with models of elastic stress transfer^{1,2}, particularly the persistence of hanging-wall seismicity³ that included two events with magnitudes greater than 5. Here we show that this sequence may have been driven by a fluid pressure pulse generated from the coseismic release of a known deep source⁴ of trapped high-pressure carbon dioxide (CO₂). We find a strong correlation between the high-pressure front and the aftershock hypocentres over a two-week period, using precise hypocentre locations⁵ and a simple model of nonlinear diffusion. The triggering amplitude (10–20 MPa) of the pressure pulse overwhelms the typical (0.1–0.2 MPa) range from stress changes in the usual stress triggering models^{1,6}. We propose that aftershocks of large earthquakes in such geologic environments may be driven by the coseismic release of trapped, high-pressure fluids propagating through damaged zones created by the mainshock. This may provide a link between earthquakes, aftershocks, crust/mantle degassing and earthquake-triggered large-scale fluid flow.

The 1997 Umbria–Marche seismic sequence in the Northern Apennines, Italy (Fig. 1a), occurred on shallow-dipping (~40°) normal faults, migrating from north-northwest (NNW) to south-southeast (SSE) along an area of about 40 km in length and 15 km in width. Structural geology studies showed that the earthquakes nucleated close to the deeper portion of a thrust, and that the faults were not optimally oriented relative to the regional stress field⁷. Non-optimally oriented faults become seismically active either because of lower friction coefficients or the presence of fluid pressures in excess of hydrostatic^{8–11}.

A geologic cross-section integrating surface geology with seismic reflection profiles^{12,13} (Fig. 1b) shows that the first two mainshocks nucleated in the Triassic evaporites (made up of alternating anhydrites and dolomites). All earthquakes of moment magnitude $M_w > 5$ nucleated in the evaporites, the same lithologic unit where CO₂ at near-lithostatic pressure was encountered in the San Donato borehole at a depth of 4.8 km about 50 km northwest of Colfiorito⁴. The tectonic environment of the Northern Apennines is suitable for trapping high-pressure fluids derived from CO₂ mantle degassing^{14,15}, and in particular, the Rasiglia spring in the epicentral region shows an area-averaged deep CO₂ production rate¹² of approximately $6 \times 10^5 \text{ mol m}^{-2} \text{ yr}^{-1}$.

We show that the driving mechanism for this earthquake and aftershock sequence is the coseismic release and propagation of the trapped high-pressure source into the overlying carbonates at hydrostatic pore pressure. Coseismic fracturing of the seal separating these two distinct pressure states initiates the rapid propagation of a pressure pulse along the newly formed, highly permeable fault zone and into the adjacent damage zone. The newly fractured crust provides high-permeability channels to propagate the pulse

and trigger seismicity by significantly reducing the effective normal stress acting on incipient slip planes. In addition, the earthquakes and aftershocks themselves create new fractures, resulting in a large-scale permeability structure that increases significantly as the sequence evolves. This is shown by recasting the epicentres (Fig. 1a) as the approximate slipped area of the sequence (Fig. 2a), demonstrating how the entire region evolves to a complex system of fractures that provide conduits for propagating the pressure pulse.

There are two regions of interest for this study. The first is the largest event of the entire sequence (event 3), and its associated

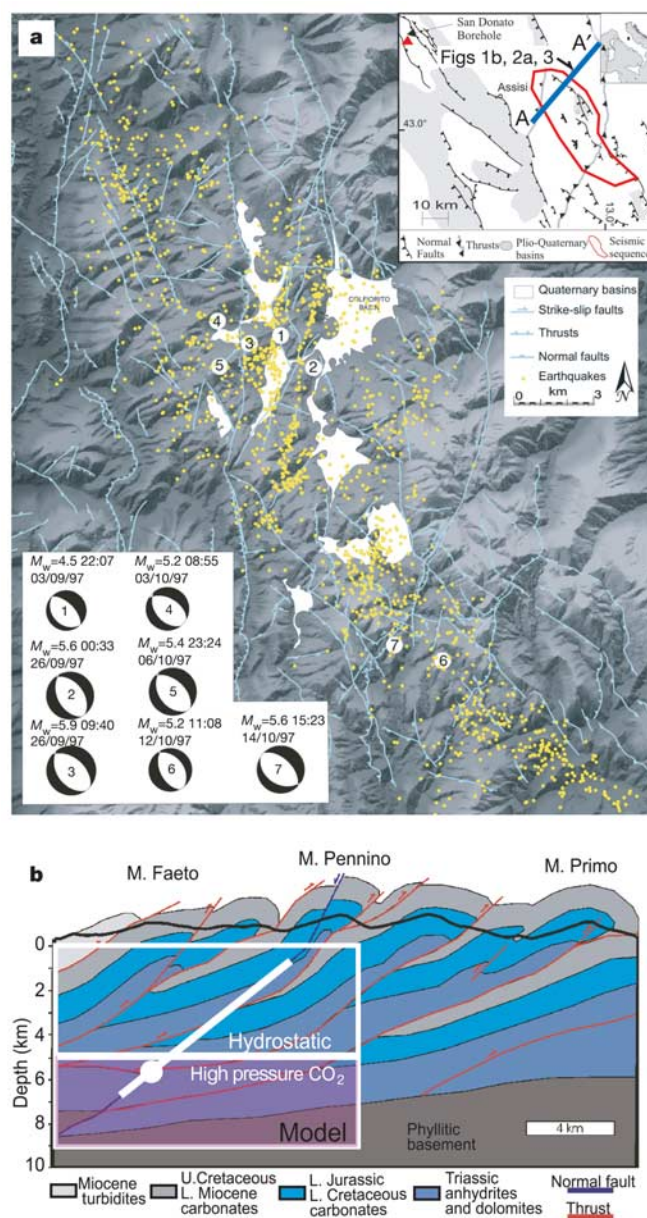


Figure 1 Geologic setting of the study area. **a**, Seismicity and major structures of the Colfiorito region of the Northern Apennines, Italy. The region consists of a complex pattern of thrusts, folds and normal faults reflecting two main tectonic phases: a Miocene–Pliocene compressional phase forming east-northeast (ENE)–verging thrusts and folds; and a superimposed upper Pliocene–Quaternary extensional phase forming basins bounded by NNW–SSE–trending normal faults^{12,13}. Near-lithostatic pore pressure (CO₂) measured in the San Donato borehole (see inset) was encountered in the evaporites and just below the seal of a sub-horizontal thrust. All $M_w > 5$ earthquakes nucleated in the evaporites. **b**, Geologic cross-section calibrated from geology and seismic profiles¹³, with the simplified model shown superposed (see also Fig. 3). U., Upper; L., Lower.

aftershocks (Fig. 2a). We can approximate and model this sequence by projecting the aftershocks onto a two-dimensional profile (A–A') that also corresponds to the geologic cross-section in Fig. 1b. The second region of interest is the sequence propagating to the SSE, where the observed propagation velocity (Fig. 2b) can be used to estimate the structural permeability of the system, and to show supporting evidence for the high permeability used in the model. Figure 2b shows that this sequence propagated at a relatively constant velocity of about 1 km d^{-1} , similar to the velocity of CO_2 -driven seismicity inferred from a swarm system¹⁶, and for induced seismicity in a deep borehole¹⁷. A fluid-pressure-induced sequence with a constant propagation velocity is consistent with the wave-like solutions of a pressure pulse found for flow problems where permeability is a strong nonlinear function of the effective normal stress $\bar{\sigma}_n$ acting on the fracture^{18,19}. Assuming that the observed

propagation velocity reflects the velocity of a wave-like pulse, we can roughly estimate the large-scale permeability of the system using the relationship¹⁸ $V = k\gamma/(\eta\phi)$, where V is the pulse velocity, k is the intrinsic permeability, γ is the weight density difference between the rock and fluid, η is the viscosity, and ϕ is the porosity. Taking $\gamma = 1.7 \text{ MPa m}^{-3}$, $\eta = 10^{-3} \text{ Pa s}$, $\phi = 0.05$, and using the observed V , then $k = 4 \times 10^{-11} \text{ m}^2$. This is a very large permeability compared to the 10^{-16} m^2 inferred for crustal permeability in tectonically stable environments²⁰, but lower than the permeability inferred for the Dobi extensional earthquake sequence in Central Afar²¹, and consistent with permeability measurements on rough fractures in granite and marble at low effective normal stress²².

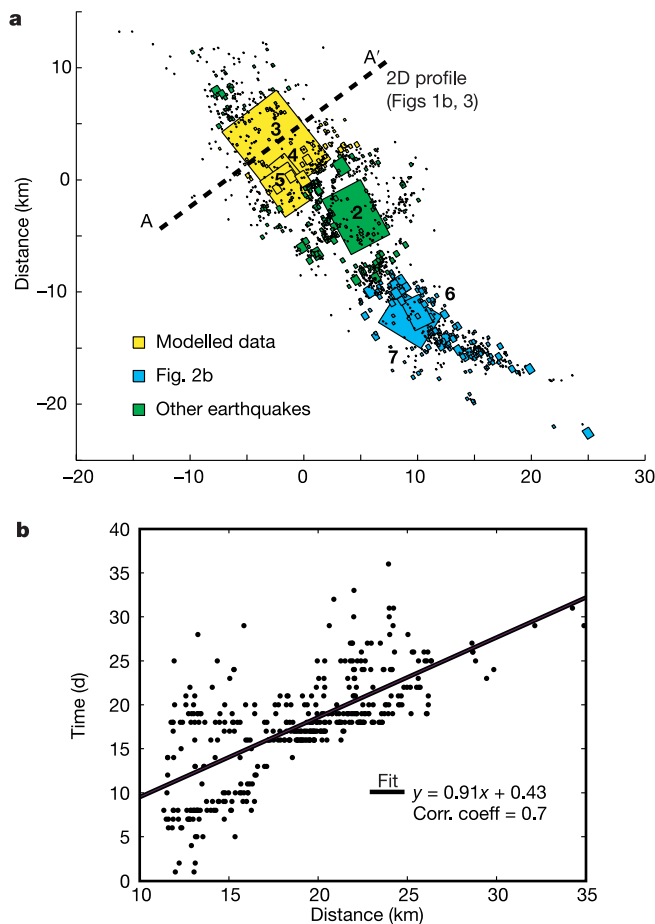


Figure 2 Map view of the seismicity and the rate of propagation. **a**, The seismicity in Fig. 1a recast as the approximate area fractured in the earthquake or aftershock to illustrate how the earthquake or aftershock themselves drastically alter the structural permeability of the system. Patches show the slipped area using the relation $M = G\Delta\sigma\bar{u}$, where M is the scalar seismic moment, G is the shear modulus (30 GPa), $\Delta\sigma$ is the stress drop (assumed to be 1.5 MPa), and slip $\bar{u}$ is calculated from $\Delta\sigma = (2G/\pi)(\bar{u}/W)$. The events are colour-coded to show the events (yellow) compared to the model results, and include all events with $M_w > 2.4$ between the hypocentre of event 3 and about 7 km to the NW. Section A–A' corresponds to the cross-section in Fig. 1b, and represents an approximate two-dimensional profile onto which the hypocentres of events shown in yellow are projected. Events shown in blue are plotted in **b** as the distance from the hypocentre of event 3 versus time to estimate the structural permeability of the system. It is the length of the vector between hypocentres, and therefore includes both the horizontal and up-dip migration of the sequence. The linear correlation implies that the propagation velocity is faster than the $t^{1/2}$ diffusion timescale, and a least-squares fit shows that this sequence propagates at $\sim 1 \text{ km d}^{-1}$ (that is, $\sim 10^{-2} \text{ m s}^{-1}$).

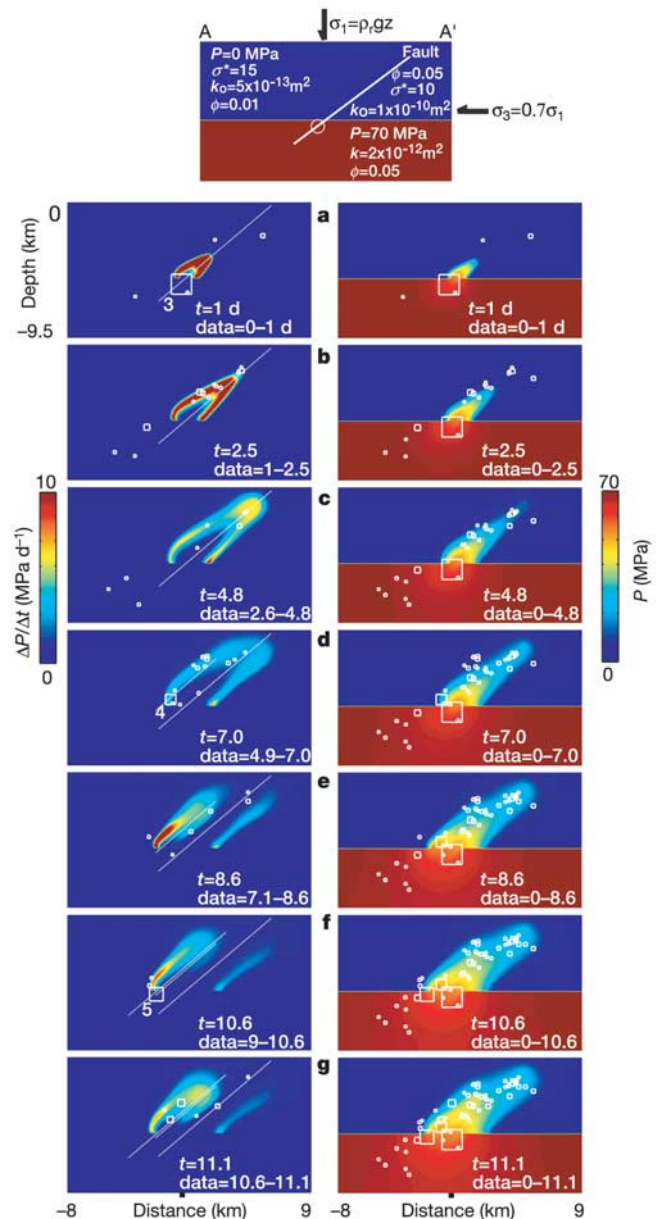


Figure 3 Comparison of model results with initial conditions (top) to the hypocentres of aftershocks (bottom) shown in yellow in Fig. 2a. **a–g**, Model results plotted as the rate of pore pressure increase to highlight propagation of the pressure front (left column), and the corresponding evolution of the entire fluid pressure field (right column). The left column compares the evolution of the pore pressure front to the aftershocks occurring during the times indicated. The overall fluid pressure field is superposed with the cumulative aftershock catalogue. The largest event in the sequence (event 3) and subsequent large aftershocks in the hanging wall (events 4 and 5) are indicated in **a**, **d** and **e**.

Modelling the three-dimensional flow field is beyond the scope of this Letter, but we can simplify the analysis by considering the two-dimensional profile perpendicular to the strike of event 3. All aftershocks with $M_w > 2.4$ in the volume surrounding event 3 are projected onto the two-dimensional profile (A–A') for comparison with model results. We model this sequence by numerically solving a nonlinear diffusion equation with an effective-stress-dependent permeability (see Methods). In the initial conditions, an impermeable seal separates the upper region at hydrostatic fluid pressure from the high-pressure source region where the initial fluid pressure was taken at 85% of lithostatic⁴ (for example, 70 MPa) at the upper boundary of the source region. A hydrostatic pressure gradient is imposed below the seal, corresponding to no-flow conditions before the earthquake. Flow initiates at $t = 0$, when a model fault and damage zone approximately 400 m in width (field observations show damage zones from 200 to 600 m) cuts the high-pressure region and extends to about 1 km below the surface. This simulates the coseismic fracturing of the pore pressure seal. The sudden communication between the high-pressure source and low-pressure surroundings initiates a pressure pulse that propagates along the fault and into the hanging and footwalls.

The evolution of the propagating pressure front and the fluid pressure field (Fig. 3) are superposed with the hypocentres of aftershocks for the period indicated in each panel. The steep front is a consequence of the wave-like solution for pressure-dependent permeability, in contrast to a more diffuse front resulting from linear diffusion. At early times (Fig. 3a, b), the pulse propagates rapidly up the fault zone, leaving in its wake a slower-moving pulse into the matrix material of the hanging and footwalls. The faster propagation into the hanging wall relative to the footwall is a result of the decreasing σ_n (thus increasing permeability) as the pulse

propagates to shallower depths. As the pulse reaches the hydrostatic boundary condition imposed at the surface, the pressure is reduced but the pulse continues to propagate. The high-pressure front arrives and triggers event 4 (Fig. 3d), which we model by introducing a second fault (with the same properties as the original fault) extending into the source region (dotted line). Because this new high-permeability fault extends into the source region, the system is recharged and generates additional aftershocks (Fig. 3e). The subsequent pulse propagates to the location of event 5 (Fig. 3f), triggering additional seismicity (Fig. 3g). Aftershocks for events 4 and 5 correlate with the rapid propagation of a pulse into the already highly pressurized footwalls of these events.

The evolution of the total fluid pressure in the system (right column of Fig. 3) shows that the data are matched in space and through time, and follow the structure of the evolving fluid pressure field. The aftershocks in the high-pressure source region appear to correlate with contours of reduced fluid pressure. In this case, the mechanism of triggering is probably the transition from aseismic slip at high pore pressures to seismic slip as fluid pressure is reduced²³, or alternatively, a complex source region.

Our model presents an alternative interpretation of the physical processes controlling earthquake triggering in the neighbourhood of the causative fault. Several studies have shown correlations with stressing rate changes^{6,24}, static stress transfer¹, or poro-elastic effects^{25,26}. These models rely on extremely small stress changes (~ 0.1 MPa) and therefore have not unequivocally demonstrated that simple static stress changes or poro-elastic effects are the dominant mechanism of earthquake triggering or driving aftershocks. The change in Coulomb failure stress (ΔCFS) is defined as $\Delta CFS = \Delta\tau + \mu(\Delta\sigma_n + \Delta P_f)$, where $\Delta\tau$ and $\Delta\sigma_n$ are the shear and normal stress changes (positive in extension), and ΔP_f is the change in pore pressure. Attempts to relate this earthquake sequence to ΔCFS from shear stress changes failed to explain this sequence², particularly for the persistence of aftershocks in the hanging wall³. Figure 4a shows the aftershock data with the ΔCFS (for stress changes only) for event 3, and Fig. 4b compares the same aftershock data with the calculated pressure field.

In most ΔCFS formulations, the focus is primarily on changes in τ and σ_n , and it was found to be difficult to properly include poro-elastic effects²⁷. For event 3, ΔCFS from shear and normal stress changes are on the order of a few tenths of an MPa. Significantly, our results show that the effect from a 10–20 MPa direct fluid pressure loading (for example, ΔP_f) overwhelms static stress transfer.

The structural, seismic and post-seismic response of this sequence support a scenario where high-pressure CO₂ infiltrated the incipient seismic fault before the large earthquakes, followed by a large-scale change in the hydraulic properties of the system. The coseismic fracture generated a high-amplitude pressure pulse initiating at the high-pressure/low-pressure boundary, propagating into the damaged region caused by the mainshock. The increased fluid pressure triggered subsequent earthquakes and aftershocks by significantly reducing the effective normal stress. The results also suggest that the aftershocks in regions of increasing pore pressure occur along contours of constant $\bar{\sigma}_n$, implying that earthquakes occurred at the same shear stress (assuming a constant friction coefficient) independently of depth. As the effect on ΔCFS due only to pore pressure changes is orders of magnitude greater than the contribution of elastic stress transfer, we propose that this mechanism dominates some triggering phenomena and aftershock sequences. Although this sequence was driven by CO₂ out-gassing, the processes of fracture and coseismic hydraulic property changes are general, suggesting this is an important general mechanism of aftershock generation. That is, earthquakes provide the trigger to hydraulically connect the upper crust at hydrostatic pore pressure with the lower crust at near-lithostatic pore pressure. The subsequent flow will be fast, high-pressured, and will propagate readily into the new fractures created by the main event. □

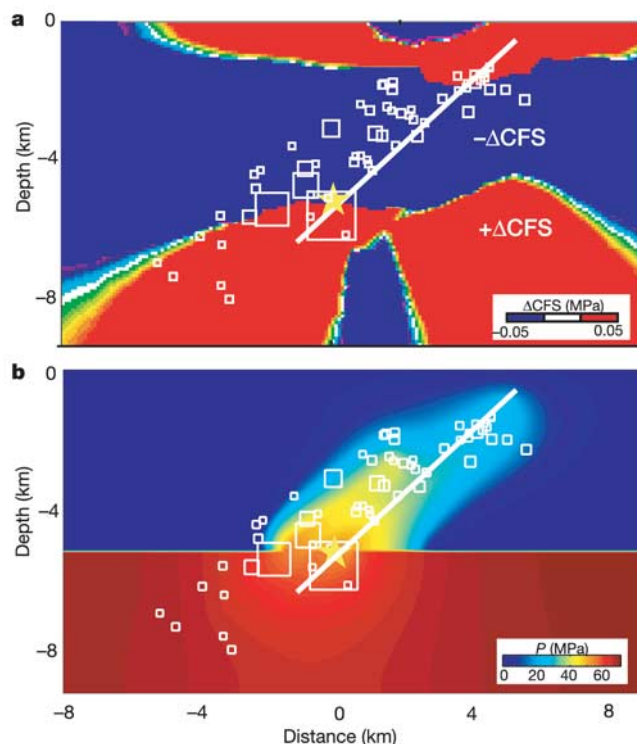


Figure 4 Comparison of aftershock data to stress changes in the ΔCFS formulation and pore pressure changes. **a**, There is no correlation between positive or negative ΔCFS regions and the aftershock locations. In contrast, **b**, the same aftershock data compared to the calculated fluid pressure state after 11 days, shows a very strong correlation with the entire aftershock sequence (see also Fig. 3.)

Methods

We adopt the model of Rice¹⁸ where permeability is a strongly decreasing function of effective normal stress, $k = f(\sigma_n)$. Specifically, $k = k_0 \exp(-\sigma_n/\sigma^*)$, where k_0 is the permeability at zero effective stress, and σ^* is a constant with lower values of σ^* corresponding to highly cracked rocks. Using this form for the permeability, we solve the diffusion equation with a spatially variable permeability²⁸:

$$\frac{\partial P}{\partial t} = \frac{1}{\phi(\beta_f + \beta_\phi)} \left[\nabla \cdot \frac{k_0 \exp(-\frac{\sigma_n}{\sigma^*})}{\eta} \nabla P + \dot{P}(P, T) \right] \quad (1)$$

where P is the fluid pressure above hydrostatic, β_f and β_ϕ are the fluid and pore (crack) compressibility, and $\dot{P}$ is a source term. The source term is assumed to be zero here, but is included in equation (1) to show that the pressure dependence of the dehydration (or de-carbonization) kinetics could provide an additional direct fluid source from coseismic fluid pressure reductions²⁹. The effective normal stress used in equation (1) and acting on fault planes is calculated as³⁰:

$$\sigma_n = \frac{\sigma_1 + \sigma_3 - 2P_f}{2} + \frac{\sigma_1 - \sigma_3}{2} \cos 2\theta \quad (2)$$

where σ_1 and σ_3 are the maximum and minimum principal stress, P_f is the total fluid pressure (for example, $P + \rho_w g z$), θ is the dip angle, ρ_w is the density of water, g is the acceleration of gravity, and z is the depth. We take $\theta = 40^\circ$ (determined from the earthquake focal mechanisms), σ_1 as the weight of the overburden (for example, $\rho_r g z$), where ρ_r is the rock density, and we assume $\sigma_3 = 0.7\sigma_1$ to reflect this extension tectonic environment.

We solve equation (1) with an implicit finite difference scheme, using the simplified model geometry and initial conditions shown at the top of Fig. 3. A no-flow boundary condition is imposed on all boundaries except the upper surface, where a constant head (for example, hydrostatic pore pressure) boundary condition is imposed. We use crack compressibility $\beta_\phi = 10^{-8} \text{ MPa}^{-1}$, fluid compressibility $\beta_f = 10^{-10} \text{ MPa}^{-1}$, and a temperature-dependent viscosity for water (assuming a temperature gradient of $25^\circ \text{C km}^{-1}$). We assume that the flow properties of supercritical CO_2 (the phase of CO_2 at the source depth) are the same as for water because CO_2 at this P - T condition is ten times more compressible than water, but it is of the same order less viscous, resulting in similar flow properties. Note that the model can be made much more complicated by considering two-phase flow, dual porosity, anisotropic permeability, and other complexities. However, we find this simple model sufficient to show a very strong correlation between the calculated pressure field and the precise locations of aftershock hypocentres.

Received 13 June; accepted 25 November 2003; doi:10.1038/nature02251.

- Stein, R. S. The role of stress transfer in earthquake triggering. *Nature* **402**, 605–609 (1999).
- Cocco, M., Nostro, C. & Ekström, G. Static stress changes and fault interaction during the 1997 Umbria-Marche earthquake sequence. *J. Seismol.* **4**, 501–516 (2000).
- Chiaraluce, L., Ellsworth, W. L., Chiarabba, C. & Cocco, M. Imaging the complexity of an active normal fault system; the 1997 Colfiorito Central Italy case study. *J. Geophys. Res.* **B 6**, 10.1029/2002JB002166 (2003).
- Chiodini, R. & Cioni, G. Gas geobarometry for hydrothermal systems and its application to various Italian geothermal areas. *Appl. Geochem.* **4**, 564–572 (1989).
- Waldhauser, F. & Ellsworth, W. L. A double-difference earthquake location algorithm: Method and application to the Northern Hayward Fault, California. *Bull. Seismol. Soc. Am.* **90**, 1353–1368 (2000).
- Toda, S., Stein, R. & Saglya, T. Evidence from the AD 2000 Izu islands earthquake swarm that stressing rate governs seismicity. *Nature* **419**, 58–61 (2002).
- Collettini, C. Structural permeability control on the post-seismic fluid discharge and aftershocks triggering: hypotheses for the 1997 Colfiorito earthquakes. *Boll. Soc. Geol. It.* **1**, 873–880 (2002).
- Sibson, R. H. Fault-valve behavior and the hydrostatic-lithostatic fluid pressure interface. *Earth Sci. Rev.* **32**, 141–144 (1992).
- Cox, S. F. Faulting processes at high fluid pressures: An example of fault valve behavior from the Wattle Gully Fault, Victoria, Australia. *J. Geophys. Res.* **100**, 12841–12859 (1995).
- Miller, S. A., Nur, A. & Olgaard, D. L. Earthquakes as a coupled shear stress-high pore pressure dynamical system. *Geophys. Res. Lett.* **23**, 197–200 (1996).
- Streit, J. E. & Cox, S. F. Fluid pressures at hypocenters of moderate to large earthquakes. *J. Geophys. Res.* **106**, 2235–2243 (2001).
- Reuter, K. J., Giese, P. & Closs, H. Lithosphere split in the descending plate; observations from the Northern Apennines. *Tectonophysics* **64**, T1–T9 (1980).
- Mirabella, F. & Pucci, S. Integration of geological and geophysical data along two sections crossing the region of the 1997–98 Umbria-Marche earthquake sequence. *Boll. Soc. Geol. It.* **1**, 891–900 (2002).
- Chiodini, G., Frondini, F., Cardellini, C., Parello, F. & Peruzzi, L. Rate of diffuse carbon dioxide Earth degassing estimated from carbon balance of regional aquifers: The case of central Apennine, Italy. *J. Geophys. Res.* **105**, 8423–8434 (2000).
- Quattrocchi, F. In search of evidence of deep fluid discharges and pore pressure evolution in the crust to explain the seismicity style of the Umbria-Marche, 1997–1998 seismic sequence (Central Italy). *Ann. Geophys.* **42**, 609–636 (1999).
- Bräuer, K., Kämpf, H., Strauch, G. & Weise, S. M. Isotopic evidence ($^3\text{He}/^4\text{He}$, $^{13}\text{C}/^{12}\text{C}$) of fluid-triggered intraplate seismicity. *J. Geophys. Res.* **108**, 10.1029/2002JB002077 (2003).
- Baisch, S. & Harjes, H.-P. A model for fluid-injection-induced seismicity at the KTB, Germany. *Geophys. J. Int.* **152**, 160–170 (2003).
- Rice, J. R. In *Fault Mechanics and Transport Properties of Rock* (eds Evans, B. & Wong, T.-f.) 476–503 (Academic, San Diego, 1992).
- Revil, A. & Cathles, L. M. III Fluid transport by solitary waves along growing faults. A field example from the South Eugene Island Basin, Gulf of Mexico. *Earth Planet. Sci. Lett.* **202**, 321–335 (2002).
- Manning, C. E. & Ingebritsen, S. E. Permeability of the continental crust: Implications of geothermal data and metamorphic systems. *Rev. Geophys.* **37**, 127–150 (1999).
- Noir, J., Jacques, E., Békri, S., Adler, P. M. & King, G. C. P. Fluid flow triggered migration of events in the 1989 Dobi earthquake sequence of Central Afar. *Geophys. Res. Lett.* **24**, 2335–2338 (1997).
- Lee, H. S. & Cho, T. F. Hydraulic characteristics of rough fractures in linear flow under normal and shear load. *Rock Mech. Rock Eng.* **35**, 299–318 (2002).

- Segall, P. & Rice, J. R. Dilatancy, compaction, and slip instability of a fluid infiltrated fault. *J. Geophys. Res.* **100**, 22155–22171 (1995).
- Dieterich, J. A constitutive law for rate of earthquake production and its application to earthquake clustering. *J. Geophys. Res.* **99**, 2601–2618 (1994).
- Nur, A. & Booker, J. R. Aftershocks caused by pore fluid flow? *Science* **175**, 885–887 (1972).
- Bosl, W. J. & Nur, A. Aftershocks and pore fluid diffusion following the 1992 Landers earthquake. *J. Geophys. Res.* **107**, 10.1029/2001JB000155 (2002).
- Cocco, M. & Rice, J. R. Pore pressure and poroelasticity in Coulomb analysis of earthquake interactions. *J. Geophys. Res.* **107**, 10.1029/2000JB000138 (2002); correction 10.1029/2002JB002319 (2003).
- Wong, T.-f., Ko, S.-c. & Olgaard, D. L. Generation and maintenance of pore pressure excess in a dehydrating system, 2, Theoretical analysis. *J. Geophys. Res.* **102**, 841–852 (1997).
- Miller, S. A., van der Zee, W., Olgaard, D. L. & Connolly, J. A. D. A fluid-pressure controlled feedback model of dehydration reactions: Experiments, modelling, and application to subduction zones. *Tectonophysics* **370**, 241–251 (2003).
- Jaeger, J. C. & Cook, N. G. W. *Rock Mechanics* (Chapman and Hall, London, 1979).

Acknowledgements We thank C. H. Scholz for comments, and D. Giardini, Y. Y. Podladchikov, J. A. D. Connolly, T. Kohl, K. Evans, G. Hillers and N. Deichmann for discussions.

Authors' contributions C.C., structural geology and concept development; L.C. and M.C., aftershock data; M.B., structural geology; B.J.P.K., numerical model development.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to S.A.M. (steve.miller@erdw.ethz.ch).

Changes in fisheries discard rates and seabird communities

Stephen C. Votier¹, Robert W. Furness¹, Stuart Bearhop^{1,2}, Jonathan E. Crane¹, Richard W. G. Caldow³, Paulo Catry⁴, Kenny Ensor¹, Keith C. Hamer⁵, Anne V. Hudson¹, Ellen Kalmbach⁵, Nicholas I. Klomp⁷, Simone Pfeiffer^{1,8}, Richard A. Phillips⁹, Isabel Prieto¹ & David R. Thompson¹⁰

¹Institute of Biomedical and Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK

²School of Biology and Biochemistry, Queens University, Belfast, Medical Biology Centre, 97 Lisburn Road, Belfast BT9 7BL, UK

³Centre for Ecology and Hydrology, Winfrith Technology Centre, Winfrith Newburgh, Dorchester, Dorset DT2 8ZD, UK

⁴Unidade de Investigação em Eco-Etologia, ISPA, Rua Jardim do Tabaco 44, Lisboa, Portugal

⁵Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK

⁶Department of Animal Ecology, University of Groningen, Kercklaan 30, 9751NN Haren, The Netherlands

⁷Johnstone Centre, School of Environmental and Information Sciences, Charles Sturt University, PO Box 789, Albury, New South Wales 2640, Australia

⁸Institute of Ecology, Dornburger Strasse 159, University of Jena, 07743 Jena, Germany

⁹British Antarctic Survey, High Cross, Madingley Road, Cambridge CB3 0ET, UK

¹⁰National Institute of Water and Atmospheric Research Ltd, PO Box 14-901, Kilbirnie, Wellington, New Zealand

It is clear that discards from commercial fisheries are a key food resource for many seabird species around the world^{1–8}. But predicting the response of seabird communities to changes in discard rates is problematic and requires historical data to elucidate the confounding effects of other, more 'natural' ecological processes. In the North Sea, declining stocks, changes in technical measures, changes in population structure⁹ and the establishment of a recovery programme for cod (*Gadus morhua*¹⁰) will alter the amount of fish discarded. This region also supports internationally important populations of seabirds¹¹, some of which feed extensively, but facultatively, on discards, in particular on undersized haddock (*Melanogrammus aeglefinus*)

and whiting (*Merlangius merlangus*)^{1–3}. Here we use long-term data sets from the northern North Sea to show that there is a direct link between discard availability and discard use by a generalist predator and scavenger—the great skua (*Stercorarius skua*). Reduced rates of discarding, particularly when coupled with reduced availability of small shoaling pelagic fish such as sandeel (*Ammodytes marinus*), result in an increase in predation by great skuas on other birds. This switching of prey by a facultative scavenger presents a potentially serious threat to some seabird communities.

Current fishery practices produce vast quantities of waste in the form of offal and the discarding of an estimated 25–30 million tonnes of undersized fish worldwide each year¹². There is also evidence that populations of scavenging seabirds have increased considerably where discards are plentiful^{1–2,11}. It is likely that declines in discard availability (from the closure of fisheries or changes in management policy) will affect seabird communities both directly and indirectly. Scavenging species are affected directly in terms of foraging ecology⁷, breeding biology⁸ and overwinter condition¹³. But only a few seabird species in each community feed on fishery discards, and these tend to be large seabirds that can compete effectively at fishing boats for this resource³.

Indirect effects include increased depredation of smaller seabirds by scavengers facing a shortfall in their energy budgets¹⁴. Such effects may be confounded by the fact that scavengers can also feed on alternative prey such as small, shoaling, lipid-rich species, for example, sandeels and capelin (*Mallotus villosus*)^{14–16}, that may

themselves fluctuate in abundance in association with the biomass of predatory fish and hence catch-per-unit effort and discard rates. Long-term data have enabled us to examine the effect of changes in discard rates, and discard use by seabirds—a natural experiment—on seabird communities as a whole.

The North Sea supports many important fisheries, including a large fishery operating mainly in the north that targets white fish (cod, haddock, whiting and saithe (*Pollachius virens*)) and generates substantial quantities of waste¹. Over the past 40 yr, there have been large declines in stocks of these fish in the North Sea⁹, during which time *per capita* mortality due to fishing has tended to increase. Although there have been large variations in discard rates from year to year depending on highly variable levels of recruitment (ref. 4 and see Methods), in general the volume of fish discarded has fallen less rapidly than the stocks⁹. The northern North Sea also supports large breeding populations of scavenging seabirds including northern fulmar (*Fulmarus glacialis*), northern gannet (*Morus bassanus*), great skua, lesser black-backed gull (*Larus fuscus*), herring gull (*Larus argentatus*) and great black-backed gull (*Larus marinus*)¹¹ and, although it is highly desirable to recover depleted gadoid stocks through reductions in fishery quotas or fishery closures¹⁰, an unavoidable consequence of such action will be an unprecedented reduction in food supply for scavenging seabirds.

The great skua is a top predator in marine food webs, feeding extensively on discards, as well as on sandeels and seabirds^{15,16}, which makes it an ideal species with which to evaluate the complex relationship between fisheries and seabird community structure. Great skuas experienced rapid population growth in the past century¹¹; this growth is almost certainly due to protection from persecution and is also related to increases in discard availability. Although this species is the subject of an intensive long-term study at the largest colony in the world (Foula, Shetland, UK; 60° 08' N, 02° 05' W), many standard breeding biology and population-level parameters do not readily lend themselves to analysing the importance of discard availability.

Drawing conclusions from changes in breeding numbers may be confounded because great skuas are long-lived with very stable populations from year to year, and changes in the recruitment of nonbreeders buffer the effects of environmental change¹⁷. In addition, because of the long delay between fledging and maturity, changes in breeding numbers are likely to lag behind changes in

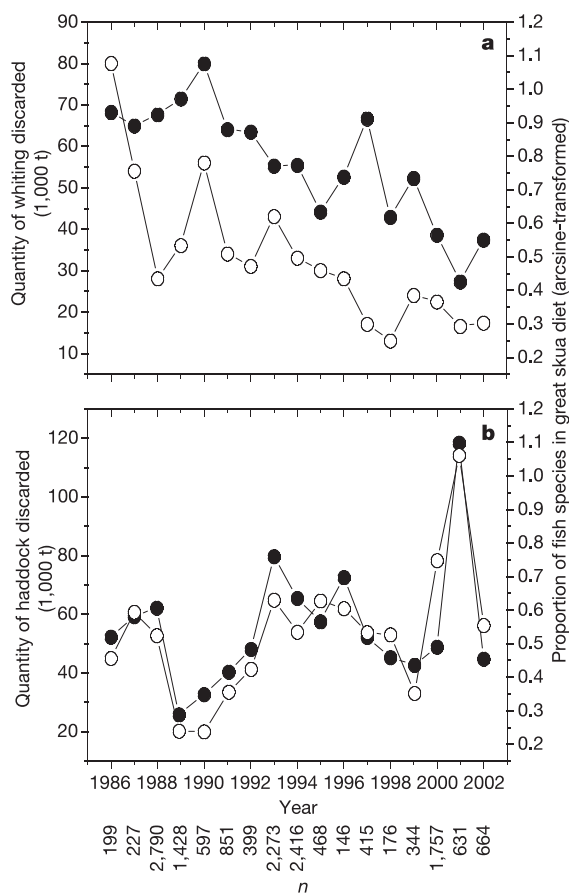


Figure 1 Composition of the diet of great skuas tracks the quantity of fishery discards. There are significant correlations between ICES estimates of discards (filled circles) and the (arcsine-transformed) percentage of these fish in the diet of great skuas (open circles) for both whiting (a) and haddock (b) for the period 1986–2002. The sample number of otoliths for each year is indicated (n).

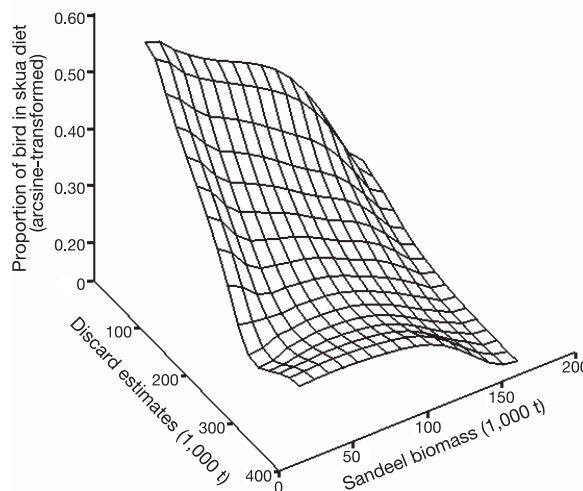


Figure 2 Declining discard and sandeel availability results in an increase in seabird predation. The relationship between the percentage of bird prey in the diet (arcsine-transformed) and both discard and sandeel availability is shown for the period 1974–1994. The relationship has been fitted by using T4253H smoothing in SPSS Advanced Statistics (Version 10.0, Chicago).

discard rates². Changes in breeding frequency and success may also occur as a consequence of other unquantified environmental factors in both the breeding and the wintering season, or may be counter-balanced by prey-switching¹⁸. But by studying the diet choice of this generalist predator over a long period of time, we can test directly the importance of discards and the use of other prey in relation to discard availability.

To investigate the use of discarded white fish species, we compared the proportion of discards in the diet of great skuas with International Council for the Exploration of the Sea (ICES) estimates of quantities of fish discarded during 1986–2002 (see Methods). These data show that there are highly significant positive correlations between discard estimates and the importance of both whiting (arcsine square-root-transformed data: correlation coefficient, $r = 0.62$, degrees of freedom, d.f. = 15, $P = 0.008$; Fig. 1a) and haddock (arcsine square-root-transformed data: $r = 0.78$, d.f. = 16, $P < 0.0001$; Fig. 1b) in the diet of great skuas.

We analysed the relationship between the (arcsine square-root-transformed) percentage of bird prey consumed by great skuas and the abundance of both of their other main food resources, white fish discards and sandeels, to test whether changes in these two principal resources influence prey-switching to other seabirds (see Methods). We found a significant, negative, additive effect of the availability of discarded white fish (Fig. 2; Wald statistic = 6.29, d.f. = 1, $P = 0.01$) and the biomass of Shetland sandeel stock (Fig. 2; Wald statistic = 19.43, d.f. = 1, $P = 0.001$) on the proportion of bird in the diet of great skuas.

Estimates of discarded whiting and haddock varied considerably during 1986–2002 (Fig. 1). Discard rates are influenced by various complex and interactive factors, including stock size, recruitment, net mesh size, gear selectivity, legal minimum landing

size, behaviour of fishermen and overall fishing intensity⁴. Therefore, the relationship between discard availability and discard use by great skuas is unlikely to be driven by unmeasured environmental factors and provides, to our knowledge, the most compelling evidence of a direct causal relationship between fishery discards and use by a top marine predator.

The declines in discard availability have coincided with long-term declines in the biomass of sandeel spawning stock, for which the lowest observed values were recorded in 2000 (ref. 9). It would seem that this has exacerbated the prey-switching tendency of great skuas and hence their likely impact on other seabird populations. Although one of the longer-term consequences of low sandeel abundance and an absence of discards may be a marked reduction in great skua breeding performance and population size, their ability to switch prey may buffer them from these consequences for some time. With relatively stable populations¹¹ and low levels of deferred breeding¹⁹, variation in the percentage of bird prey in great skua diets will almost certainly reflect changes in predation rates on other seabirds.

There is evidence from elsewhere in Shetland that great skua predation reduced one large population of black-legged kittiwakes (*Rissa tridactyla*) by 54–85% between 1981 and 1995 (ref. 20), and negatively affected adult survival of black-legged kittiwakes on Foula²¹. Bioenergetics models indicate that a slight increase in the percentage of bird in the diet results in large numbers of other seabirds being consumed by great skuas. For example, a 5% increase in bird in the great skua diet at Foula is equivalent to the consumption of an additional 1,000 northern fulmars or 2,000 black-legged kittiwakes (Box 1). Although it is extremely difficult to quantify this impact on prey populations, this may reflect a substantial component of annual mortality (Box 1). We may there-

Box 1

Bio-energetics modelling to quantify seabird predation by great skuas

By using published values of energetic requirements of great skuas, as well as population estimates, it is possible to calculate the approximate amount of food needed to support the adult great skua population on Foula. With these data, we can explore a hypothetical model in which bird prey in the diet increases by 5%. We use published estimates of the calorific content of bird meat to calculate the mass of bird required to meet this requirement. Finally, by dividing this total by the body mass values of specific prey species, we can estimate the number of birds that would need to be eaten, if this model were true. We present population estimates of adult seabird prey (breeding numbers from published surveys in 2000 (ref. 27) and estimates of nonbreeders calculated by using published life-history parameters of each species for productivity, annual survival and age of first breeding) to provide a context for this theoretical predation rate.

It has been calculated¹⁶ that adult great skuas require 2,480 kJ per bird per day for maintenance and activity (field metabolic rate). In the first 15 d of July (the period of diet sampling in this study), around 5,000 adult great skuas are present on Foula (breeders and nonbreeders combined)^{17,27}, giving an energetic requirement for this period of 1.86×10^8 kJ (that is, $15 \times 2,480 \times 5,000$). If we assume a 5% increase in birds in the diet, this would represent 9.3×10^6 kJ, which, for calorific

values of 10.9 kJ g^{-1} for whole birds, would be equivalent to 853.2 kg of birds. If this energetic demand were met by consuming fully grown seabirds of a single species, the estimates shown in Box 1 Table 1 would apply.

We cannot, however, easily quantify the demographic groups that are affected by great skuas, which has implications for assessing the impact on population levels. For species such as fulmars and puffins that fledge chicks late in the summer, predation by great skuas is almost completely on adults. For kittiwakes and guillemots, predation is on adults and fledglings, and in some colonies on chicks. Observations of prey remains indicate that predation of adult birds predominates in April to mid-July, whereas predation of fledglings forms a larger part of the diet in late July. Predation of adult seabirds will affect populations of long-lived species more rapidly than will predation of eggs and chicks²⁸; however, it is rarely possible to distinguish between pellets containing feathers from adult and those containing feathers from young birds. The effect of killing only an additional few per cent of the adult population of seabirds can be profound, given the high survival rates and low recruitment characteristic of these species, and can change a low rate of population growth into a population decline²⁸.

Box 1 Table 1 Estimates of single species needed to support a 5% increase in seabird consumption by great skuas

Species	Adult population estimate	Body mass (kg)*	Number of birds consumed†	Adult population consumed (%)
Northern fulmar	64,800	0.8	1,067	1.6
Black-legged kittiwake	6,400	0.41	2,094	32.8
Common guillemot	54,000	0.86	992	1.8
Atlantic puffin	55,000	0.4	2,133	3.9

*Data from ref. 29.

†Equivalent to 853.2 kg of bird prey/body mass.

fore predict that other seabirds such as northern fulmar, European storm-petrel (*Hydrobates pelagicus*), Leach's storm-petrel (*Oceanodroma leucorhoa*), common guillemot (*Uria aalge*) and Atlantic puffin (*Fratercula arctica*), which are already killed in large numbers by great skuas¹⁶, might experience marked population declines before the skuas do themselves, because the skuas are able to switch from feeding on discards or sandeels.

Current numbers of seabirds may be artificially maintained in the northwestern North Sea owing to depleted stocks of piscivorous fish, which reduces competition for small schooling fish, and to the provision of discards and offal by fisheries²². In some areas, however, rates of predation on seabirds by great skuas have already reached levels that seem to be unsustainable^{18,20}. It is unlikely that this predator simply has a beneficial role in maintaining healthy populations by removing sick or weak individuals.

Although it would not be appropriate to maintain current rates of discarding for the sake of seabirds, further drastic cuts in white fish catches in the North Sea will exacerbate the problem of great skua predation in the short term and may lead to marked reductions in seabird populations. It is clear, however, that sandeel abundance, in addition to white fish discard rates, influences prey-switching in this top marine predator (Fig. 2). Thus, measures to conserve stocks of sandeels should be maintained and perhaps given even greater priority to avert the scenario suggested here that is likely to occur if the discarding of white fish were to cease. In the longer term it is unclear whether the North Sea seabird community will return to a more 'natural' former state, or whether the effects of diet switching and predation will cause it to lurch to a completely different composition. □

Methods

Annual variations in discard availability

Data on the quantities of white fish discarded in the northwestern North Sea are provided by ICES^{9,23}. Since 1960, there have been several step changes in technical measures such as net mesh size, minimum landing size and single-species quotas. These have interacted with changes in the behaviour of fishermen (such as the amount of selective upgrading by discarding marketable fish), resulting in discard rates of whiting falling faster than catches of this fish^{9,23}. The proportion (by mass) of whiting caught in the North Sea that is discarded has decreased significantly in the past 40 yr ($r = -0.54$, d.f. = 40, $P < 0.05$); by contrast, the proportion of haddock discarded has not changed^{9,23}. Despite the changes in the discard rate of whiting, the annual catch (tonnes) explains 91.5% of the variance in whiting discarded annually from 1960 to 2001; the relationship is best described by an exponential relationship: discard mass = $11.783e^{0.0114(\text{catch mass})}$. For haddock, the annual catch (tonnes) explains 61% of the variance in the quantity of haddock discarded annually from 1960 to 2001; the relationship is best described by a linear regression: discard mass = $0.3538(\text{catch mass}) + 10.95$.

Diet composition

Indigestible material regurgitated as pellets provides a reliable assessment of diet composition in great skuas²⁴. Pellets were collected annually from club sites of nonbreeding birds on Foula, Shetland, in the first half of July (the peak period of the breeding season of skuas) during 1974–2002. Pellets of breeders were collected in only 6 yr; however, the proportion of bird in the pellets from nonbreeders' club sites correlates with that in the pellets from breeders' territories sampled in the same years (arcsine square-root-transformed $r = 0.83$, d.f. = 5, $P = 0.02$). Each pellet was identified to the lowest possible taxon and removed to prevent recounting. Any fish otoliths were stored dry before being measured and were identified by a guidebook²⁵.

To investigate the importance of discarded white fish species in skua diets, we applied linear regression to compare the proportion of haddock and whiting otoliths in skua diets (arcsine square-root-transformed) and ICES discard estimates of these two species. Haddock and whiting are demersal species, living at depths where great skuas, and other species that the skuas may kleptoparasitise such as northern gannet, are unable to catch them. In addition, the size of fish in the diet of great skuas (calculated from the length of otoliths²⁵) is predominantly just below the commercial size limit (haddock: mean 27.6 cm, range 21.1–36.9 cm, $n = 4,739$; whiting: mean 30.1 cm, range 18.5–40.2 cm, $n = 8,032$), indicating that these are discards²⁶.

Using residual maximum likelihood models, we constructed a mixed model with (arcsine square-root-transformed) percentage of bird pellets (total number of pellets: 1974, $n = 100$; 1975, $n = 558$; 1976, $n = 483$; 1977, $n = 100$; 1978, $n = 100$; 1979, $n = 100$; 1980, $n = 100$; 1981, $n = 100$; 1982, $n = 100$; 1983, $n = 305$; 1984, $n = 100$; 1985, $n = 100$; 1986, $n = 200$; 1987, $n = 98$; 1988, $n = 598$; 1989, $n = 796$; 1990, $n = 187$; 1991, $n = 100$; 1992, $n = 694$; 1993, $n = 117$; 1994, $n = 199$) in the diet as the response variable, and tonnes of sandeels in the Shetland stock using virtual population analysis (VPA) from 1974 to 1994 (ref. 9; the period for which these data are available) and tonnes of haddock and whiting (combined) discarded in the northwestern North Sea (ICES

subarea IV) as predictor covariates. All interactions were nonsignificant and were therefore removed from the model (Table 1).

To correct for problems of temporal autocorrelation and independence of data points, we entered covariates of all independent variables into the model for values at year _{t} /($n+1$)/year _{t} . None of these terms was significant in the final model.

Received 12 December 2003; accepted 5 January 2004; doi:10.1038/nature02315.

- Garthe, S., Camphuysen, C. J. & Furness, R. W. Amounts of discards by commercial fisheries and their significance as food for seabirds in the North Sea. *Mar. Ecol. Prog. Ser.* **136**, 1–11 (1996).
- Furness, R. W. Impacts of fisheries on seabird communities. *Sci. Mar.* **67** (suppl. 2), 33–45 (2003).
- Hudson, A. V. & Furness, R. W. The behaviour of seabirds foraging at fishing boats around Shetland. *Ibis* **131**, 225–237 (1989).
- Stratoudakis, Y., Fryer, R. J., Cook, R. M. & Pierce, G. J. Fish discarded from Scottish demersal vessels: Estimators of total discards and annual estimates for targeted gadoids. *ICES J. Mar. Sci.* **56**, 592–605 (1999).
- Bearhop, S. *et al.* Annual variation in great skua diets: the importance of commercial fisheries and predation on seabirds revealed by combining dietary analyses. *Condor* **103**, 802–809 (2001).
- Thompson, K. R. Quantitative analysis of the use of discards from squid trawlers by black-browed albatrosses *Diomedea melanophris* in the vicinity of the Falkland Islands. *Ibis* **134**, 11–21 (1992).
- Arcos, J. M. & Oro, D. Changes in foraging range of Audouin's gulls *Larus audouinii* in relation to a trawler moratorium in the western Mediterranean. *Colonial Waterbirds* **19**, 128–131 (1996).
- Oro, D. The effects of trawler discard availability on the egg-laying and the breeding success of the lesser black-backed gull *Larus fuscus* in the western Mediterranean. *Mar. Ecol. Prog. Ser.* **132**, 43–46 (1996).
- ICES. *Report of the ICES Advisory Committee on Fishery Management, 2002*. Cooperative Research Report No. 255 (ICES, Copenhagen, Denmark, 2002).
- ICES. *Advises Zero Catches of Cod and Other Fish Stocks*. ICES Press Release 20 October 2003; at <http://www.ices.dk/aboutus/pressrelease/ACFMAutumn2003.pdf> (2003).
- Lloyd, C. S., Tasker, M. L. & Partridge, K. *The Status of Seabirds in Britain and Ireland* (Poyser, London, 1991).
- Alverson, D. L., Freeberg, M. H., Murawski, S. A. & Pope, J. G. *A Global Assessment of Fisheries Bycatch and Discards*. FAO Fisheries Technical Paper No. 339 (1994).
- Hüppop, O. & Wurm, S. Effects of winter fishery activity on resting numbers, food and body condition of large gulls *Larus argentatus* and *L. marinus* in the south-eastern North Sea. *Mar. Ecol. Prog. Ser.* **194**, 241–247 (2000).
- Russell, J. O. & Montevecchi, W. A. Predation on adult puffins *Fratercula arctica* by great black-backed gulls *Larus marinus* at a Newfoundland colony. *Ibis* **138**, 791–794 (1996).
- Hamer, K. C., Furness, R. W. & Caldow, R. W. G. The effects of changes in food availability on the breeding ecology of great skuas *Catharacta skua* in Shetland. *J. Zool. Lond.* **223**, 175–188 (1991).
- Phillips, R. A., Thompson, D. R. & Hamer, K. C. The impact of great skua predation on seabird populations at St Kilda: a bioenergetics model. *J. Appl. Ecol.* **36**, 218–232 (1999).
- Klomp, N. I. & Furness, R. W. Non-breeders as a buffer against environmental stress: declines in numbers of great skuas on Foula, Shetland, and prediction of future recruitment. *J. Appl. Ecol.* **29**, 341–348 (1992).
- Regehr, H. M. & Montevecchi, W. A. Interactive effects of food shortage and predation on breeding failure of black-legged kittiwakes: effects of fisheries activities and implications for indicator species. *Mar. Ecol. Prog. Ser.* **155**, 249–260 (1996).
- Catry, P., Phillips, R. A., Hamer, K. C., Ratcliffe, N. & Furness, R. W. The incidence of nonbreeding by adult great skuas and parasitic jaegers from Foula, Shetland. *Condor* **100**, 448–455 (1998).
- Heubeck, M., Mellor, R. M., Harvey, P. V., Mainwood, A. R. & Riddington, R. Estimating the population size and rate of decline in kittiwakes *Rissa tridactyla* breeding in Shetland 1981–97. *Bird Study* **46**, 48–61 (1999).
- Oro, D. & Furness, R. W. Influences of food availability and predation on survival of kittiwakes. *Ecology* **83**, 2516–2528 (2002).
- Furness, R. W. Management implications of interactions between fisheries and sandeel-dependent seabirds and seals in the North Sea. *ICES J. Mar. Sci.* **59**, 261–269 (2002).
- ICES. *Report of the Working Group on the assessment of demersal stocks in the North Sea and Skagerrak*. ICES CM 1998/Assess: 7 (ICES, Copenhagen, Denmark, 1998).
- Votier, S. C., Bearhop, S., MacCormick, A., Ratcliffe, N. & Furness, R. W. Assessing the diet of great skuas *Catharacta skua* using five different techniques. *Polar Biol.* **26**, 20–26 (2003).
- Härkönen, T. *Guide to the Otoliths of the Bony Fishes of the Northeast Atlantic* (Danbiu ApS, Hellerup, Denmark, 1986).
- Furness, R. W. & Hislop, J. R. G. Diets and feeding ecology of Great Skuas *Catharacta skua* during the breeding season in Shetland. *J. Zool. Lond.* **195**, 1–23 (1981).
- Harvey, P., Gear, S., Swale, J. & Upton, A. in *Shetland Bird Report 2000* (ed. Pennington, M.) 99–102 (Shetland Bird Club, Shetland, UK, 2001).
- Croxall, J. P. & Rothers, P. in *Bird Population Studies: Relevance to Conservation and Management* (eds Perrins, C. M., Lebreton, J.-D. & Hiron, G. J. M.) 272–296 (Oxford Univ. Press, Oxford, UK, 1991).
- Snow, D. W. & Perrins, C. M. *The Birds of the Western Palearctic, Concise Edition* Vol. 1 (Oxford Univ. Press, Oxford, UK, 1998).

Acknowledgements We thank S. Humphries, N. Ratcliffe, A. Kelly and S. Waldron for assistance and discussion; and all those who helped over the years to collect prey remains at Foula. We thank the Fishery Research Service, Aberdeen, for providing Shetland sandeel VPA data. The research was supported by the Shetland Oil Terminal Environmental Advisory Group (SOTEAG) and by the European Commission contract 'Discbird'. S.B. is funded by a Natural Environment Research Council (NERC) postdoctoral fellowship. P.C. is funded by the Fundação para Ciência e a Tecnologia (FCT), Portugal.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to S.C.V. (s.votier@bio.gla.ac.uk) or R.W.F. (rfurness@bio.gla.ac.uk).

Soil biota and exotic plant invasion

Ragan M. Callaway, Giles C. Thelen, Alex Rodriguez & William E. Holben

Division of Biological Sciences, The University of Montana, Missoula, Montana 59812, USA

Invasive plants are an economic problem and a threat to the conservation of natural systems. Escape from natural enemies might contribute to successful invasion¹, with most work emphasizing the role of insect herbivores^{2–4}; however, microbial pathogens are attracting increased attention⁵. Soil biota in some invaded ecosystems may promote 'exotic' invasion^{6–9}, and plant–soil feedback processes are also important. Thus, relatively rare species native to North America consistently demonstrate negative feedbacks with soil microbes that promote biological diversity¹⁰, whereas abundant exotic and native species demonstrate positive feedbacks that reduce biological diversity¹⁰. Here we report that soil microbes from the home range of the invasive exotic plant *Centaurea maculosa* L. have stronger inhibitory effects on its growth than soil microbes from where the weed has invaded in North America. *Centaurea* and soil microbes participate in different plant–soil feedback processes at home compared with outside *Centaurea*'s home range. In native European soils, *Centaurea* cultivates soil biota with increasingly negative effects on the weed's growth, possibly leading to its control. But in soils from North America, *Centaurea* cultivates soil biota with increasingly positive effects on itself, which may contribute to the success of this exotic species in North America.

Soil microbes have profound negative and beneficial effects on plants through pathogenic effects, root–fungus mutualisms and by driving the nutrient cycles on which plants depend^{5,11–17}. These effects, and the reciprocal effects of plants on soil microbes, contribute to two contrasting dynamic feedback interactions between plants and the microbial communities that develop around their roots^{18,19}. Positive feedbacks occur when plant species accumulate microbes near their roots that have beneficial effects on the plants that cultivate them, such as mycorrhizal fungi and nitrogen fixers. Positive feedbacks are thought to lead to a loss of local community diversity^{18,19}. Negative feedbacks occur when plant species accumulate pathogenic microbes in their rhizospheres, creating conditions that are increasingly hostile to the plants that cultivate the pathogens^{10,20,21}. Negative feedbacks are thought to enhance community diversity by increasing species turnover rates.

Here we show that the direct effects of soil biota and the feedback loops that develop between soil biota and an invasive plant depend on the biogeographical source of the microbes. We compared the effects on *C. maculosa* growth of soil microbial communities collected from four populations of *C. maculosa* in its native range in western Europe to the effects of soil microbes collected from six populations in the northwestern United States where *C. maculosa* has invaded. On average, sterilization of European soils caused a 166% increase in the total biomass of *C. maculosa* compared with a 24% increase when North American soils were sterilized (Fig. 1). Soil sterilization can cause nutrient flushes; however, all pots were well fertilized and the effects of sterilization were consistent for multiple *C. maculosa* populations sampled over wide areas in Europe and North America, suggesting that the sterilization effect is not simply a peculiarity of a given site. However, the effect of soil sterilization varied among populations within continents. Depending on the *C. maculosa* population from which the soil microbial community was sampled and whether or not soil was sampled from *C. maculosa* or grass rhizospheres, sterilization of European soils improved *C. maculosa* growth from as little as 31% to over 900%, whereas the effects of sterilizing North American soils ranged from a

24% decrease in *C. maculosa* growth (suggesting a positive effect of microbes) to a 148% increase (Supplementary Information). The stronger suppressive effects of European soil biota lend experimental support to earlier demonstrations of much higher fungal and viral infection on plant species in their home ranges than in invaded ranges⁵, and indicate that *C. maculosa* in North America have escaped the controlling effects of soil biota.

We further examined biogeographical differences in plant–soil microbe relationships by using soils from the *C. maculosa* population in the Central Massif in France and soils from Missoula, Montana, in a feedback experiment. We chose these because they represented the extreme cases: strong positive sterilization effects (inhibitory soil biota) on soils from the native range in the Central Massif, and weak negative sterilization effects (beneficial soil biota) on soils from the invaded range in Missoula (Supplementary Fig. 1). The microbial community in the soil from France was 'pre-cultured'^{18–21} by planting either *C. maculosa* or *Festuca ovina*, a small perennial bunchgrass that is native to Eurasia, in pots with the original soils and allowing the plants to interact with the soil community for 3 months. The microbial community in the soil from Montana was pre-cultured by planting *C. maculosa* or *Festuca idahoensis*, a bunchgrass similar to *F. ovina* but native to western North America, in pots with the original soils. The pre-cultured soils with their microbial communities were then used to inoculate media in which we grew *C. maculosa* alone or in competition with one of the two grass species.

As observed in the first experiment, *C. maculosa* plants grown in Montana soils were much larger than those grown in French soils when soils were not sterilized. However, in this experiment substantial feedback loops between soil microbes and *C. maculosa* were demonstrated (Fig. 2). *Centaurea maculosa* plants grown alone in non-sterile French soil pre-cultured by conspecifics were significantly smaller than those grown in French soils pre-cultured by *F. ovina*. In contrast, *C. maculosa* planted alone in Montana soils

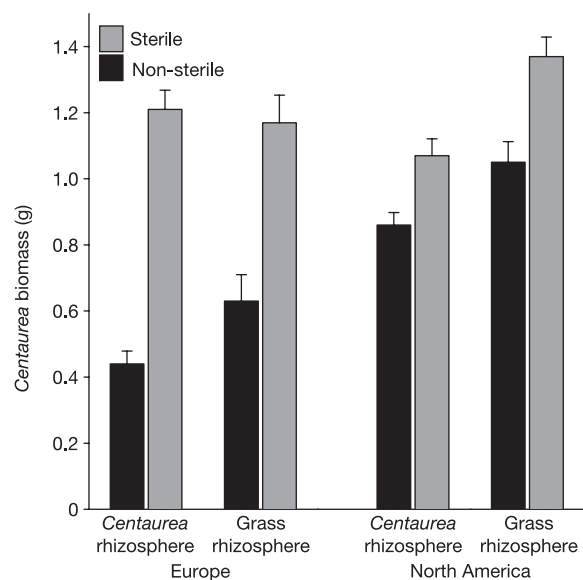


Figure 1 Total biomass of *C. maculosa* plants grown in non-sterilized and sterilized soil collected from European ($n = 4$) and North American ($n = 6$) populations of *C. maculosa*. Soils were collected from the rhizospheres of neighbouring grasses and from the rhizospheres of *C. maculosa*. In a 4-way analysis of variance (ANOVA) (region of origin, rhizosphere and sterilization as main effects, and population nested within region of origin) $F_{\text{origin}} = 40.18$, degrees of freedom (d.f.) = 1,297, $P < 0.001$; $F_{\text{pop}} = 21.81$, d.f. = 5,297, $P < 0.001$; $F_{\text{rhiz}} = 21.12$, d.f. = 1,297, $P < 0.001$; $F_{\text{sterilization}} = 110.87$, d.f. = 1,297, $P < 0.001$; $F_{\text{origin} \times \text{sterilization}} = 29.35$, d.f. = 1,297, $P < 0.001$; $F_{\text{origin} \times \text{pop} \times \text{rhiz}} = 16.2$, d.f. = 2,314, $P < 0.001$.

pre-cultured by conspecifics were significantly larger than in Montana soils pre-cultured by *F. idahoensis*. Sterilization of the soils eliminated these feedbacks. The negative effects of soil microbes are often enhanced in competitive environments²², but we also found results almost identical to those in Fig. 2 for *C. maculosa* grown in competition with bunchgrasses in pre-cultured French or Montana soils (data not shown). Considered together, the results of the feedback experiments suggest that *C. maculosa* is able to modify the microbial community in invaded soils to its advantage. In contrast, *C. maculosa* is inhibited by a negative feedback in French soils, probably due to the accumulation of pathogens and potentially also due to adaptation of inhibitory microbial populations to antimicrobial compounds produced by *C. maculosa*²³. However, the generality of these results should be interpreted with caution because of our choice to experiment with soils from sites with the maximal contrasting sterilization results.

In a field experiment designed to investigate plant–soil feedbacks in more realistic conditions in North America, we found that *C. maculosa* growing in locations that had been occupied by other *C. maculosa* plants for 3 years were over twice as large as other *C. maculosa* plants growing in sites occupied by the native grass, *Pseudoroegneria spicata* (0.67 ± 0.12 versus 0.30 ± 0.07 g, *t*-test, $P = 0.023$).

To examine descriptively the pre-culturing effects of different plant species on the composition of soil microbial communities from different regions, we conducted a comparative analysis of soil microbial community structure in the Central Massif (France) and

Missoula (North America) soils using denaturing gradient gel electrophoresis of partial 16S and 18S rRNA gene fragments amplified by polymerase chain reaction (PCR) using the generally conserved primers 536fC and 907r. The results suggested that *C. maculosa* cultured European and North American microbial communities in different ways, and that the effects of *F. idahoensis* differed from that of *C. maculosa* (Supplementary Fig. 2). However, on the basis of these results alone we cannot determine whether *C. maculosa* stimulates or suppresses specific microbial populations.

Pathogen–plant relationships are often quite host-specific or vary substantially in their relative effects on different plant species^{10,20,24}. If plants and pathogens co-evolve locally it would be expected that feedback between a plant species and soil microbes from its native range will be negative, and that exotic invaders may escape more pathogens than they acquire in their new habitat⁵. Mutualistic microbes can have host-specific associations and functions^{25,26}, but in contrast to the host-specific tendency of pathogenic microbes, many arbuscular mycorrhizal fungi tend to infect a broad range of hosts^{25,27} (although there is evidence for specialization in the mutualistic benefits for the plant²⁸), making it possible for invaders to use the native mycorrhizae of a new region. Therefore, the feedback of soil microbes from the invaded range of an exotic weed to the weed itself is likely to be neutral or positive because of the potential for the invader to accumulate mutualistic fungi in the absence of host-specific soil pathogens.

Our results support Klironomos' hypothesis¹⁰ that some exotic invaders have escaped control by local soil pathogens, and may benefit from soil microbes in invaded regions. However, conclusive proof would require precise matching of seeds from exotic plant populations with their co-occurring soil microbes. We intentionally used seeds of *C. maculosa* and grasses from sources that did not match any specific population in an effort to eliminate any local bias that might confound continental comparisons. Until local exotic plant populations and their specific soil microbial communities are tested together, our understanding of escape from local pathogens remains incomplete.

Many mechanisms are involved in the expansion of exotic plant species. Our results provide comparative biogeographical experiments testing mechanisms for invasive success, and show that a switch from negative plant–soil microbial feedback in native habitats to positive plant–soil feedbacks in invaded habitats may contribute to the expansion of one of the world's worst invaders. □

Methods

Experiment 1

We collected soils from four sites in Europe and six sites in North America to assess biogeographical differences in the effects of soil microbes on *C. maculosa*, a native of Europe and an invasive weed in North America. The European sites were near Ales in the Central Massif, France ($44^{\circ} 51' 0''$ N; $0^{\circ} 52' 0''$ E), near Briançon, France ($44^{\circ} 53' 60''$ N; $6^{\circ} 39' 0''$ E), near Gap, France ($44^{\circ} 34' 0''$ N; $6^{\circ} 4' 60''$ E), and at Pontamafrey, Italy ($45^{\circ} 43' 60''$ N; $7^{\circ} 9' 60''$ E). The North American sites were Missoula, Montana ($46^{\circ} 52' 23''$ N; $113^{\circ} 59' 45''$ W), the National Bison Range, Montana ($47^{\circ} 20' 12''$ N; $114^{\circ} 13' 46''$ W), Montana State University Experimental Ranch, Montana ($47^{\circ} 03' 30''$ N; $113^{\circ} 12' 42''$ W), Sapphire Mountains, Montana ($46^{\circ} 22' 52''$ N; $113^{\circ} 56' 56''$ W), The Flathead Valley, Montana ($47^{\circ} 33' 30''$ N; $113^{\circ} 42' 60''$ W) and Clearwater, Montana ($47^{\circ} 0' 34''$ N; $113^{\circ} 23' 53''$ W). Soils were collected and immediately subjected to slow air-drying to mimic drying conditions that would occur during natural drought. After drying, one-half of the soils were treated by triple autoclaving on three successive days to kill soil microbes. Soil was mixed with sand (15:85 soil:sand mixture) within 4 weeks of collection into 525-cm³ pots. *Centaurea* seeds from a single North American population for which soils were not used (Bozeman, Montana) were planted into all pots and grown for 140 days while being fertilized once every 2 weeks with 100 ml of 0.25 strength Hoaglands solution and watered every 2 days. After this period *Centaurea* plants were harvested, dried at 60 °C and weighed. Owing to an error in the autoclaving of grass rhizospheres in the Central Massif and Missoula, the sample sizes were reduced to 2–4 for each treatment site combination.

Experiments 2 and 3

Plant–soil microbe feedbacks were compared using soil from one *C. maculosa* population in Europe (Central Massif) and one in North America (Missoula). We chose to contrast these soils because they represented the extreme cases in each region. We

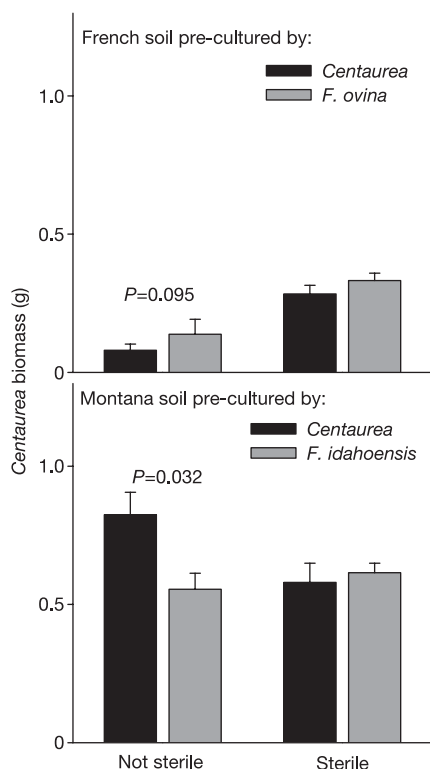


Figure 2 Total biomass of *C. maculosa* plants grown alone in European soil (Central Massif population) and North American soil (Missoula population) that had been pre-cultured by either *C. maculosa* or a *Festuca* species native to the place of soil origin. Plants were grown in soils either sterilized or not sterilized after pre-culturing. In a 3-way ANOVA (origin, species used for pre-culturing and sterilization as main effects) $F_{\text{origin}} = 159.7$, d.f. = 179, $P < 0.001$; $F_{\text{culture} \times \text{species}} = 0.30$, d.f. = 179, $P = 0.593$; $F_{\text{sterilization}} = 1.28$, d.f. = 179, $P = 0.260$; $F_{\text{origin} \times \text{culture} \times \text{species}} = 7.64$, d.f. = 179, $P = 0.007$; $F_{\text{origin} \times \text{sterilization}} = 13.99$, d.f. = 179, $P < 0.001$; $F_{\text{culture} \times \text{species} \times \text{sterilization}} = 6.03$, d.f. = 179, $P < 0.017$. P-values shown above paired bars indicate a significant difference in pre-culturing effects for those treatments.

pre-cultured^{1,2,10,20,21} the microbial communities from each site by growing either *C. maculosa* from the same region used in experiment 1, or a *Festuca* bunchgrass native to the region of soil origin, in pure soil. The two *Festuca* species (*idahoensis* and *ovina*) were chosen because they are similar in size and appearance and co-occur naturally with *C. maculosa* in their respective native lands. The use of a single *Festuca* species might have biased the relationship between the plant and microbes of a different continent in ways that might differ from a plant and microbes from the same region. Five plants were grown for 110 days in each of 40 4-litre pots: $n = 10$ for Montana soil with *C. maculosa*; Montana soil with *F. idahoensis*; French soil with *C. maculosa*; and French soil with *F. ovina*. After the 110-day pre-culturing period, half of the soil in each pot was triple-autoclaved on three successive days to kill the microbial community, and then the non-sterilized and sterilized soil from each 4-litre pot was used to inoculate two sterile 15:85 soil:sand mixtures and two non-sterile 15:85 soil:sand mixtures in 525 cm³ pots. We used the soil in the 525 cm³ pots for two experiments: one in which *C. maculosa* was grown alone in sterile and non-sterile soil, and one in which *C. maculosa* was grown with *F. idahoensis* as a competitor. In the no-competition experiment, *Centaurea* seeds were planted in all pots and grown for 91 days during which they were fertilized and watered as in experiment 1. In the competition experiment, *F. idahoensis*, which grows more slowly than *C. maculosa*, was planted first, and after 14 days *C. maculosa* seeds were planted in half of the pots and grown for 91 days. At the end of the experiments all plants were harvested, dried at 60 °C and weighed.

Experiment 4

We compared the post-removal effects of *C. maculosa* on the growth of conspecifics to those of the native grass *Pseudoroegneria spicata*. Ten *C. maculosa* and ten *P. spicata* were planted in random locations outdoors in the Deittart Experimental Gardens in April 2001. The site historically supported native grasslands. These plants were grown until 10 August 2003, when they were harvested aboveground. One *C. maculosa* was grown at each of the 20 sites from 10 August to 14 September 2003 when they were harvested aboveground, dried at 60 °C and weighed.

Received 30 September 2003; accepted 6 January 2004; doi:10.1038/nature02322.

- Crawley, M. J. *Natural Enemies: the Population Biology of Predators, Parasites, and Diseases* (Blackwell Science, Oxford, UK, 1992).
- Louda, S. M., Pemberton, R. W., Johnson, M. T. & Follet, P. A. Non-target effects: the Achilles heel of biological control? *Annu. Rev. Entomol.* **48**, 365–396 (2003).
- Goeden, R. D. & Andres, L. A. *Handbook of Biological Control* (eds Bellows, T. S. & Fisher, T. W.) Ch. 34 (Academic, New York, 1999).
- McEvoy, P. & Coombs, E. M. Biological control of plant invaders: regional patterns, field experiments, and structured population models. *Ecol. Appl.* **9**, 387–401 (1999).
- Mitchell, C. G. & Power, A. G. Release of invasive plants from fungal and viral pathogens. *Nature* **421**, 625–627 (2003).
- Richardson, D. M., Allsopp, N., D'Antonio, C. M., Milton, S. J. & Rejmánek, M. Plant invasions—the role of mutualisms. *Biol. Rev.* **75**, 65–93 (2000).
- Simberloff, D. & Von Holle, B. Positive interactions of nonindigenous species: invasional meltdown? *Biol. Invas.* **1**, 21–31 (1999).
- Marler, M. M., Zabinski, C. A. & Callaway, R. M. Mycorrhizae indirectly enhance competitive effects of an invasive forb on a native bunchgrass. *Ecology* **80**, 1180–1186 (1999).
- Callaway, R. M., Newingham, B., Zabinski, C. A. & Mahall, B. E. Compensatory growth and competitive ability of an invasive weed are enhanced by soil fungi and native neighbors. *Ecol. Lett.* **4**, 1–5 (2001).
- Kironomos, J. Feedback with soil biota contributes to plant rarity and invasiveness in communities. *Nature* **417**, 67–70 (2002).
- Packer, A. & Clay, K. Soil pathogens and spatial patterns of seedling mortality in a temperate tree. *Nature* **440**, 278–281 (2000).
- Burdon, J. J. The structure of pathogen populations in natural plant communities. *Annu. Rev. Phytopathol.* **31**, 305–348 (1993).
- Van der Putten, W. H., Vet, L. E. M., Harvey, J. A. & Wäckers, F. L. Linking above and belowground multitrophic interactions of plants, herbivores, pathogens and their antagonists. *Trends Ecol. Evol.* **16**, 547–551 (2001).
- Holah, J. C. & Alexander, H. M. Soil pathogenic fungi have the potential to affect the coexistence of two tall-grass prairie species. *J. Ecol.* **87**, 598–606 (1999).
- Brundrett, M. Mycorrhizas in natural ecosystems. *Adv. Ecol. Res.* **21**, 171–313 (1991).
- Crowley, D. E., Yant, Y. C., Reid, C. P. P. & Szanislo, P. J. Mechanisms of iron acquisition from siderophores by microorganisms and plants. *Plant Soil* **130**, 179–198 (1991).
- Newsham, K. K., Fitter, A. H. & Watkinson, A. R. Root pathogenic and arbuscular mycorrhizal fungi determine fecundity of asymptomatic plants in the field. *J. Ecol.* **82**, 805–814 (1994).
- Bever, J. D., Westover, K. M. & Antonovics, J. Incorporating the soil community into plant population dynamics: the utility of the feedback approach. *J. Ecol.* **85**, 561–573 (1997).
- Bever, J. D. Negative feedback within a mutualism: host-specific growth of mycorrhizal fungi reduces plant benefit. *Proc. R. Soc. Lond. B* **269**, 2595–2601 (2002).
- Van der Putten, W. H., Van Dijk, C. & Peters, B. A. M. Plant-specific soil-borne diseases contribute to succession in foredune vegetation. *Nature* **362**, 53–56 (1993).
- Bever, J. D. Feedback between plants and their soil communities in an old field community. *Ecology* **75**, 1965–1977 (1994).
- Van der Putten, W. H. & Peters, B. A. M. How soil-borne pathogens may affect plant competition. *Ecology* **78**, 1785–1795 (1997).
- Bais, H. P., Walker, T. S., Stermitz, F. R., Hufbauer, R. S. & Vivanco, L. M. Enantiomer-dependent phytotoxic and antimicrobial activity of (±)-catechin. A rhizosecreted racemic mixture from spotted knapweed. *Plant Physiol.* **128**, 1173–1177 (2002).
- Angspurger, C. K. *Pests, Pathogens, and Plant Communities* (eds Burdon, J. J. & Leather, S. R.) Ch. 3 (Blackwell Scientific, Oxford, 1990).
- Streitwolf-Engel, R., Boller, R., Weimken, A. & Sanders, I. R. Clonal growth traits of two *Prunella* species are determined by co-occurring arbuscular mycorrhizal fungi from a calcareous grassland. *J. Ecol.* **85**, 181–191 (1997).

- Young, N. R. & Mytton, L. R. The response of white clover to different strains of *Rhizobium trifolii* in hill land reseeded. *Grass Forage Sci.* **38**, 13–39 (1983).
- Eom, A. & Hartnett, D. C. Host plant species effects on arbuscular mycorrhizal fungal communities in tallgrass prairie. *Oecologia* **122**, 435–444 (2000).
- van der Heijden, M. G. A. *et al.* Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* **396**, 69–72 (1998).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank E. Corcket and R. Michalet for assistance with locating and identifying *C. maculosa* populations in Europe, and K. Feris for assistance with denaturing gradient gel electrophoresis data analysis. Our research on soil microbes and plant invasion is supported by NSF, USDA, the Andrew W. Mellon Foundation and The University of Montana.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.M.C. (ray.callaway@mso.umt.edu).

Organization of genetic variation in individuals of arbuscular mycorrhizal fungi

Teresa E. Pawlowska & John W. Taylor

Department of Plant and Microbial Biology, University of California, Berkeley, California 94720-3102, USA

Arbuscular mycorrhizal (AM) fungi (Glomeromycota) are thought to be the oldest group of asexual multicellular organisms. They colonize the roots of most land plants, where they facilitate mineral uptake from the soil in exchange for plant-assimilated carbon¹. Cells of AM fungi contain hundreds of nuclei. Unusual polymorphism of ribosomal DNA observed in individual spores of AM fungi inspired a hypothesis that heterokaryosis—that is, the coexistence of many dissimilar nuclei in cells—occurs throughout the AM fungal life history^{2,3}. Here we report a genetic approach to test the hypothesis of heterokaryosis in AM fungi. Our study of the transmission of polymorphic genetic markers in natural isolates of *Glomus etunicatum*, coupled with direct amplification of rDNA from microdissected nuclei by polymerase chain reaction, supports the alternative hypothesis of homokaryosis, in which nuclei populating AM fungal individuals are genetically uniform. Intraspore rDNA polymorphism contained in each nucleus signals a relaxation of concerted evolution⁴, a recombination-driven process that is responsible for homogenizing rDNA repeats⁵. Polyploid organization of glomeromycotan genomes could accommodate intranuclear rDNA polymorphism and buffer these apparently asexual organisms against the effects of accumulating mutations.

Molecular phylogeny⁶ and the fossil record⁷ date Glomeromycota to the Ordovician period and indicate that AM-like fungi participated in the transition of early plants to the terrestrial habitat⁸. With no evidence of sexual reproduction, these fungi may represent an ancient asexual lineage that is much older than the asexual bdelloid rotifers⁹. The glomeromycotan reproductive mode, and their reputed position of ancient asexuals, could be verified by methods of phylogenetics and population genetics; however, this work cannot be accomplished until the genetic variation observed in individuals of AM fungi is explained.

Two models can explain the organization of genetic variation in AM fungi: first, the diverse rDNA variants may be distributed among different nuclei (heterokaryosis, Fig. 1a); or second, all

rDNA variants may be contained in each nucleus (homokaryosis, Fig. 1b). The first model implies that the fungus maintains a stable assemblage of several different genomes during its life cycle and transmits them from generation to generation^{2,3}. Such a multi-genomic configuration of individual genetic variation would set Glomeromycota apart from other organisms; however, some of the data supporting heterokaryosis have been disputed^{10,11}. The alternative hypothesis of homokaryotic organization of genetic variation entails a relaxation of concerted evolution of rDNA in the lineage. Concerted evolution is a universal phenomenon responsible for the homogenization of rDNA repeats within a repeat array, among the arrays dispersed in an individual genome, and throughout a recombining population⁵.

To distinguish between these two models, we devised three approaches. First, we investigated maintenance of genetic variation in an AM fungus from one generation to the next during clonal reproduction in laboratory single-spore cultures. Second, we analysed the structure of genetic variation in a natural AM fungal population. Third, we directly examined the rDNA variant distribution among nuclei in spores by amplifying rDNA by polymerase chain reaction (PCR) from individually microdissected nuclei.

To examine genetic variation as it occurs in a natural population of an AM fungus, we sampled a ubiquitous morphospecies, *G. etunicatum* (Glomerales), from an agricultural field in California that had been under cereal cultivation for 80 years. We used four randomly selected field isolates to start single-spore cultures on carrot roots transformed with *Agrobacterium rhizogenes*. A search for genetic markers that varied at the intrasporal (within an individual) level retrieved a marker with 56% amino acid sequence identity to a *Saccharomyces cerevisiae* gene encoding the catalytic subunit of DNA polymerase- α , *POL1*, which is a single-locus gene in yeast and other organisms.

The four *G. etunicatum* field isolates contained 16 variants of the *POL1*-like sequence (PLS), which separated into two distinct phylogenetic clusters, PLS1 and PLS2 (Fig. 2a). In the absence of information about the ploidy level and about the physical organization of the PLS variants in AM fungi in general and in *G. etunicatum* in particular, and to make our approach more specific, we focused on the PLS1 cluster, which contained 14 variants. To distinguish whether the PLS1 variation is a consequence of the neutral accumulation of mutations or selection, we examined differences between the numbers of nonsynonymous and synonymous nucleotide substitutions expected under neutrality, and those observed in PLS1 putative amino acid sequences. By Fisher's exact test¹², we were unable to reject the null hypothesis of the neutral evolution of PLS1.

Each of the four single-spore cultures representing the four field isolates produced many clonal progeny spores. Five progeny spores were collected from each culture, and PLS1 fragments were

independently amplified by PCR from each spore, cloned and sequenced. In each culture, all five spores analysed contained the same 13 PLS1 variants (Fig. 2b). This result is consistent with a homokaryotic model of organization of genetic variation (Fig. 1b) and would not be consistent with a heterokaryosis model (Fig. 1a) unless different nuclei, each carrying 1 of the 13 variants, were apportioned to all progeny spores.

To evaluate the probability of recovering all 13 variants in five spores of each of four cultures under the assumption of hetero-

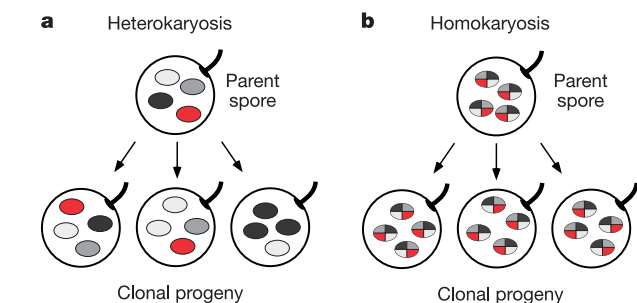
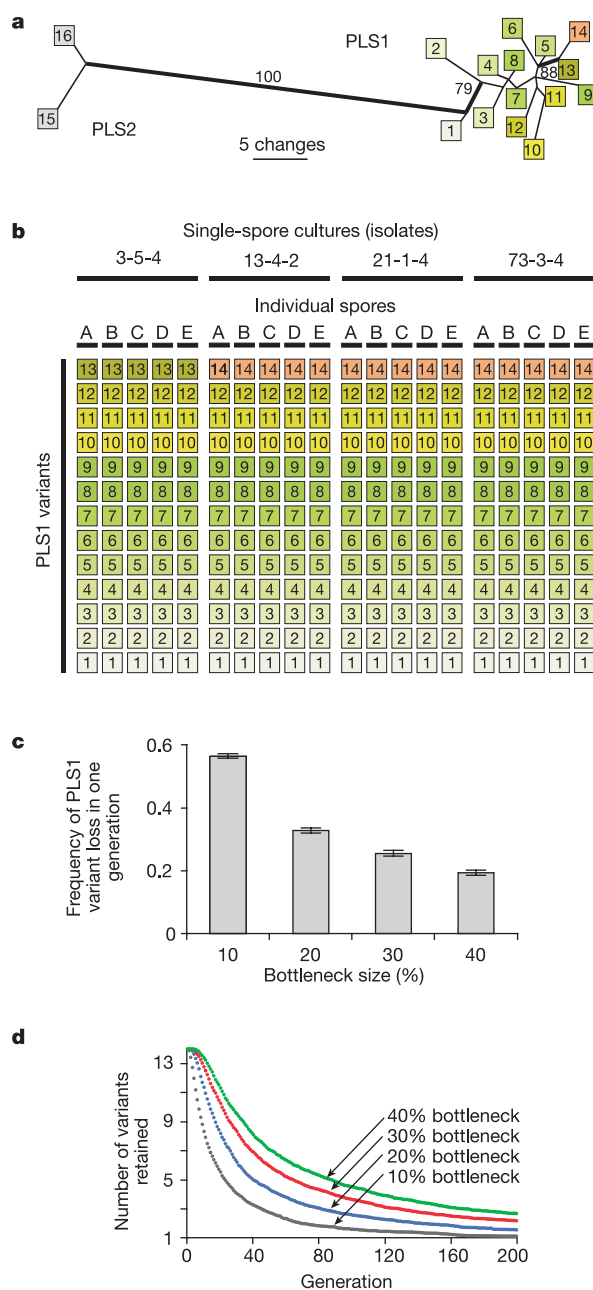


Figure 1 Sorting of variants of a polymorphic marker in single-spore cultures of *Glomus*. **a**, Sorting under the heterokaryotic model; **b**, sorting under the homokaryotic model. Variants of the polymorphic genetic marker contained in nuclei are represented by different colours.

Figure 2 PLS1 variation in *G. etunicatum* field isolates. **a**, Maximum parsimony phylogram of the PLS variants in *G. etunicatum*²⁵. Numbers beside thickened branches indicate bootstrap support values. **b**, PLS1 variants detected in natural *G. etunicatum* isolates (single-spore cultures). **c**, Expected frequency (probability) of genetically differentiated spores arising in one generation in single-spore cultures of a heterokaryotic fungus owing to loss of PLS1 variants (from the initial 13 variants) under different bottleneck schemes. Bars represent the mean $\pm$ s.e.m. expected frequencies. **d**, Summary of simulated loss of PLS1 variants over many clonal generations in a heterokaryotic fungus experiencing different bottleneck pressures.

karyosis, we modelled the transmission of a neutral polymorphic marker over one generation. To construct the model, we assumed that a spore contains 750 nuclei and that each nucleus carries 1 of the 13 PLS1 variants. We established the number of nuclei per spore by counting nuclei in several *G. etunicatum* spores. Previous work has estimated that 40% of nuclei in a spore migrate into the expanding mycelium¹³. As there are no data on the percentage of nuclei migrating from hyphae into developing spores, we used a range of possibilities (100%, 75%, 50% and 25%) in our model.

Combining the bottlenecks imposed on nuclei at spore germination and spore formation, we modelled the transmission of nuclei from parent spore to progeny spore with bottlenecks of 40%, 30%, 20% and 10%. Under heterokaryosis, our model predicted that the probability of genetic differentiation from spore to spore in one generation by loss of PLS1 variants ranged from 0.1934 (for the 40% bottleneck) to 0.5643 (for the 10% bottleneck; Fig. 2c). For four founding spores, each with five progeny spores, the probability that none of the progeny spores would lose a PLS1 variant ranged from 0.0136 to 6.1×10^{-8} , depending on the assumed bottleneck. Clearly, a model of heterokaryotic individuals is incompatible with our empirical results.

Next, we analysed the genetic variation among the four individuals representing the four isolates from the field population of *G. etunicatum*. The set of PLS1 variants was identical in three of the four individuals. These three individuals differed from the fourth individual (isolate 3-5-4) by only the replacement of variant 13 with variant 14 (Fig. 2b). To estimate the probability of finding three

individuals with identical PLS1 variants among four randomly collected individuals under the model of heterokaryosis, we simulated maintenance of genetic variation over several generations in a natural population. Our simulation began with a founder spore containing 14 types of nucleus corresponding to the 14 PLS1 variants detected in the field-isolated individuals. In the absence of any evidence for sexual reproduction, and on the basis of evidence that vegetative incompatibility mechanisms present in AM fungi prevent casual flux of nuclei among genetically distinct individuals¹⁴, we assumed strictly clonal nuclear transmission from generation to generation.

In our simulation, the number of variants of a polymorphic marker declined over time in each spore lineage with a rate dependent on the size of the bottleneck (Fig. 2d), and each lineage evolved to a different complement of variants. Under the most conservative scheme, all individuals in the field would have had identical genotypes before the maize plants, whose root systems we sampled to recover the fungus, were planted. As a result, at the time of collection the sampled isolates would be differentiated by only one generation of independent propagation. During this generation and the additional three generations mandated by our procedures between the collection of isolates and their analysis, loss of PLS1 variation could occur at the frequency estimated by our simulations (Fig. 2c).

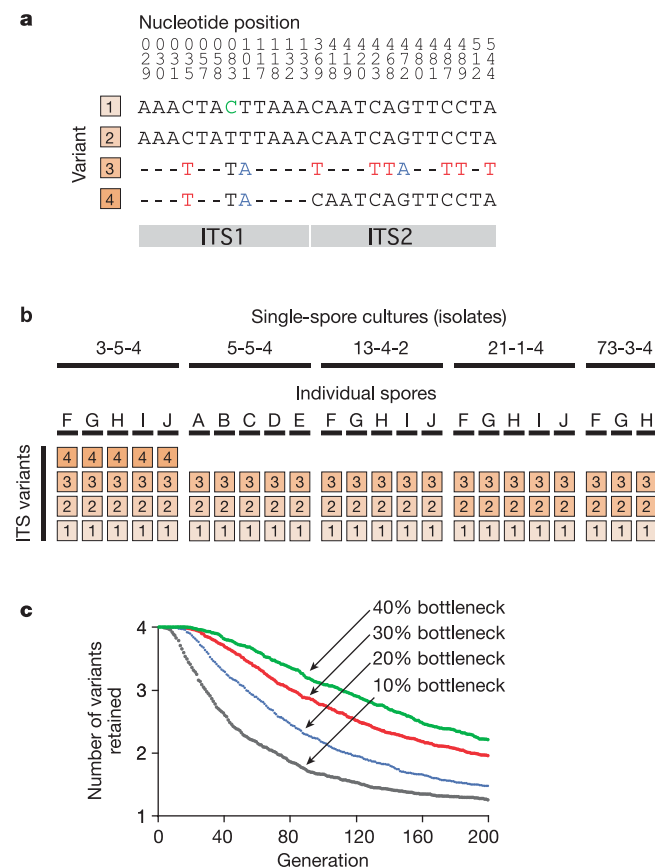


Figure 3 rDNA variation in *G. etunicatum* field isolates. **a**, Alignment of variable nucleotide positions of the ITS1–5.8S–ITS2 rDNA variants in *G. etunicatum*. **b**, ITS1–5.8S–ITS2 rDNA variants detected in natural *G. etunicatum* isolates and single-spore cultures. **c**, Summary of simulated loss of rDNA variants over many clonal generations in a heterokaryotic fungus under different bottleneck pressures.

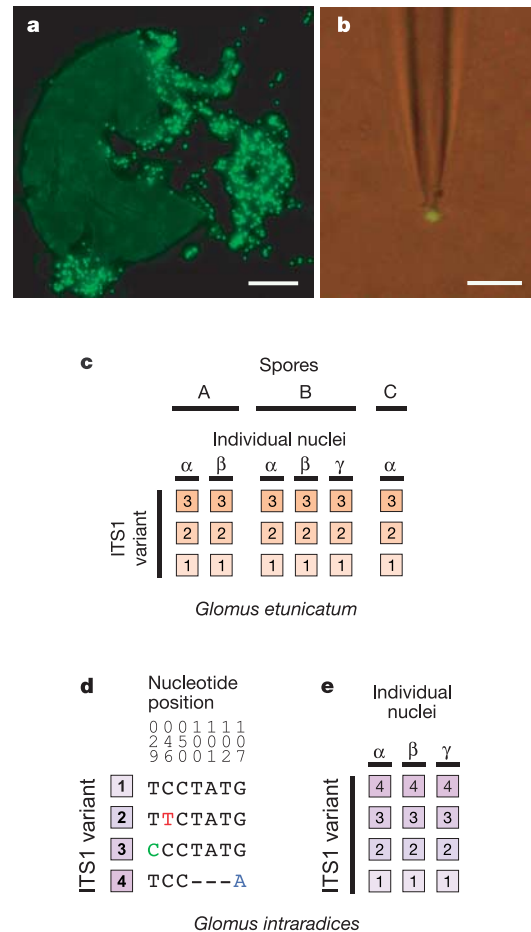


Figure 4 Distribution of rDNA variation among individual nuclei in two *Glomus* species. **a**, *G. etunicatum* spore crushed to release nuclei. Scale bar, 50 μ m. **b**, Micropipette with a nucleus attached to its tip. Scale bar, 10 μ m. **c**, ITS1 rDNA variants detected in individual nuclei of *G. etunicatum*. **d**, Alignment of variable nucleotide positions of the ITS1 rDNA variants in *G. intraradices*. **e**, ITS1 rDNA variants detected in individual nuclei of *G. intraradices*.

Consequently, the probability of recovering three isolates with identical genotypes under a model of heterokaryosis ranged from 0.0759 (40% bottleneck) to 4.7×10^{-5} (10% bottleneck). A more realistic model assumes one fungal generation per year for the 80 years since the establishment of the field (Fig. 2d) and a limited dispersal ability of AM fungal propagules¹⁵. Under this model, the probability of recovering three of four heterokaryotic isolates with identical genotypes ranged from 4.0×10^{-23} (40% bottleneck) to 2.0×10^{-87} (10% bottleneck). Thus, our empirical data are inconsistent with a model of heterokaryosis.

In addition to the analysis of PLS1 variation, we applied a similar strategy to examine the patterns of rDNA variation among and within five *G. etunicatum* field isolates using the internal transcribed spacer (ITS) region ITS1–5.8S–ITS2 as a marker. We detected four ITS1–5.8S–ITS2 variants (Fig. 3a). Variants 1 and 2 differed at a single nucleotide position (Fig. 3a). Variants 1 and 2 differed from variant 3 at ten or nine nucleotide positions, respectively, representing 8 million years of divergence (assuming independent variant evolution and a substitution rate for ITS of 1.0×10^{-9} nucleotide substitutions per site per year¹⁶). Variant 4 was best described as a recombinant between variants 2 and 3. As in the case of PLS1, clonal spores that were formed in each of the five single-spore cultures representing field isolates were genetically identical within a culture (Fig. 3b). Our simulations indicated that the rDNA variant composition should differ among the field sampled individuals if the fungus was heterokaryotic (Fig. 3c). But four of the five field isolates contained the same complement of the three ITS1–5.8S–ITS2 variants (Fig. 3b). The probability of such a sampling outcome in a heterokaryotic fungus ranged from 0.0151 (40% bottleneck) to 1.9×10^{-5} (10% bottleneck) under the assumptions described above for PLS1 with 80 generations of independent lineage propagation. Again, the empirical data do not support heterokaryosis.

To test further the hypothesis that genetic variation in AM fungi is distributed among different nuclei, we developed a technique to microdissect nuclei from individual spores of *G. etunicatum* (Fig. 4a, b). Microdissecting individual nuclei, while preserving the ability to use them as PCR templates, is technically challenging, but we were able to amplify by PCR, clone and sequence directly the ITS1 rDNA region from six individual nuclei, representing three separate spores. Each of the six nuclei contained all three of the ITS1 variants that we previously detected in the *G. etunicatum* field isolates (Fig. 4c), providing direct evidence for the homokaryosis model in which each nucleus contains all of the rDNA variation present in a spore. PCR amplification and analysis of ITS1 variants from individually microdissected nuclei of a laboratory strain of *Glomus intraradices* (Fig. 4d, e) also showed that polymorphic rDNA sequences are present in individual nuclei, confirming that homokaryosis is not the exclusive property of *G. etunicatum* but constitutes normal genetic structure in Glomerales.

The highly divergent rDNA variants present in individual nuclei of AM fungi signify a molecular pattern of relaxation in the concerted evolution of rDNA. Polymorphism of rDNA sequences may reflect an incomplete homogenization of diverse rDNA arrays that were brought together by an interspecific hybridization event¹⁷, or it may result from impediments in meiotic recombination leading to independent evolution of arrays located on homologous chromosomes¹⁸. Independent rDNA evolution may also occur in arrays distributed among nonhomologous chromosomes if these arrays do not participate in interchromosomal genetic exchanges¹⁹. The recombinant rDNA chimaeras in *G. etunicatum* indicate that recombination, which underlies the concerted evolution process, occurs in AM fungi. It remains to be seen whether this is solely mitotic recombination within an individual, or whether it also involves genetic exchanges and cryptic sexual recombination.

The rDNA heterogeneity combined with the large number of PLS variants in *G. etunicatum* may be an expression of a duplicated genome structure²⁰, which would explain the relatively large

genome size of about 10^8 base pairs estimated for Glomerales²¹. Genome polyploidization has been detected in many asexual taxa²², and it is an effective way in which to buffer the genome against the effects of accumulating mutations²³. In the absence of sexual reproduction in Glomeromycota, genome polyploidization accompanied by periodic changes in the ploidy level²⁴ might be the mechanism accounting for their long-term evolutionary persistence. The understanding that genetic variation in individuals of AM fungi is contained within each nucleus provides a foundation to address the glomeromycotan reproductive mode in nature and to examine their ancient asexual status. □

Methods

Experimental procedures

Experimental details are given in the Supplementary Information.

Computer simulations and probability estimates

To simulate clonal transmission of PLS1 variants over several generations under the heterokaryosis scheme, we populated spores with 750 nuclei, each carrying a single PLS1 variant. The founder spore contained 14 PLS1 variants distributed at random among the nuclei, and each nucleus was sampled with replacement from the 14 types of PLS1 variant. To create the *t* generation spore, nuclei were sampled with replacement from the *t* – 1 generation spore. The number of sampled nuclei corresponded to the size of a bottleneck imposed on the nuclear population. Mutations in nuclei were introduced by assigning a new variant to a random nucleus with a frequency of 2×10^{-4} mutations per generation¹⁶. The nuclei deposited in the *t* generation spore were multiplied at random to reach the count of 750. We replicated the simulation 250 times for each bottleneck size. Transmission of nuclei carrying different ITS variants over generations was simulated in a similar manner. Each nucleus was assigned a label representing one of the four different ITS1–5.8S–ITS2 rDNA variants. Mutations occurred at a rate of 3×10^{-4} per generation¹⁶. Transmission of rDNA under each bottleneck condition was replicated 300 times.

We also estimated a frequency (probability) of genetic differentiation of spores formed in single-spore cultures of a heterokaryotic fungus over one generation owing to loss of PLS1 and ITS1–5.8S–ITS2 rDNA variants. A parent spore was populated with 750 nuclei, each carrying 1 of 13 variants for PLS1, or 1 of 3 variants for rDNA. Distributions of variant frequency in the parent spores were generated in the multigeneration simulations described above (starting with 14 variants for PLS1 and 4 variants for rDNA). The number of nuclei determined by the bottleneck parameter was transferred to each progeny spore by sampling with replacement. We implemented mutations as before. Under each bottleneck scheme, 750 and 1,100 single-spore cultures, yielding 100 progeny spores each, were simulated for PLS1 and rDNA, respectively. Details of probability estimates are given in the Supplementary Information.

Received 16 September; accepted 15 December 2003; doi:10.1038/nature02290.

- Smith, S. E. & Read, D. J. *Mycorrhizal Symbiosis* (Academic, San Diego, 1997).
- Kuhn, G., Hijiri, M. & Sanders, I. R. Evidence for the evolution of multiple genomes in arbuscular mycorrhizal fungi. *Nature* **414**, 745–748 (2001).
- Bever, J. D. & Morton, J. Heritable variation and mechanisms of inheritance of spore shape within a population of *Scutellospora pellucida*, an arbuscular mycorrhizal fungus. *Am. J. Bot.* **86**, 1209–1216 (1999).
- Arnheim, N. *et al.* Molecular evidence for genetic exchanges among ribosomal genes on nonhomologous chromosomes in man and apes. *Proc. Natl Acad. Sci. USA* **77**, 7323–7327 (1980).
- Dover, G. Molecular drive: a cohesive mode of species evolution. *Nature* **299**, 111–117 (1982).
- Simon, L., Bousquet, J., Lévesque, R. C. & Lalonde, M. Origin and diversification of endomycorrhizal fungi and coincidence with vascular land plants. *Nature* **363**, 67–69 (1993).
- Redecker, D., Kodner, R. & Graham, L. E. Glomeralean fungi from the Ordovician. *Science* **289**, 1920–1921 (2000).
- Remy, W., Taylor, T. N., Hass, H. & Kerp, H. Four hundred-million-year-old vesicular arbuscular mycorrhizae. *Proc. Natl Acad. Sci. USA* **91**, 11841–11843 (1994).
- Welch, D. M. & Meselson, M. Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. *Science* **288**, 1211–1215 (2000).
- Redecker, D., Hijiri, M., Dulieu, H. & Sanders, I. R. Phylogenetic analysis of a dataset of fungal 5.8S rDNA sequences shows that highly divergent copies of internal transcribed spacers reported from *Scutellospora castanea* are of ascomycete origin. *Fungal Genet. Biol.* **28**, 238–244 (1999).
- Schüßler, A. Glomerale SSU rRNA gene diversity. *New Phytol.* **144**, 205–207 (1999).
- Zhang, J. Z., Kumar, S. & Nei, M. Small-sample tests of episodic adaptive evolution—a case study of primate lysozymes. *Mol. Biol. Evol.* **14**, 1335–1338 (1997).
- Bécard, G. & Pfeffer, P. E. Status of nuclear division in arbuscular mycorrhizal fungi during *in vitro* development. *Protoplasma* **174**, 62–68 (1993).
- Giovannetti, M. *et al.* Genetic diversity of isolates of *Glomus mosseae* from different geographic areas detected by vegetative compatibility testing and biochemical and molecular analysis. *Appl. Environ. Microbiol.* **69**, 616–624 (2003).
- Klironomos, J. N. & Moutoglou, P. Colonization of nonmycorrhizal plants by mycorrhizal neighbours as influenced by the collembolan, *Folsomia candida*. *Biol. Fertil. Soils* **29**, 277–281 (1999).
- Kasuga, T., White, T. J. & Taylor, J. Estimation of nucleotide substitution rates in eukaryotic fungi. *Mol. Biol. Evol.* **19**, 2318–2324 (2002).
- Wendel, J. E., Schnabel, A. & Seelan, T. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). *Proc. Natl Acad. Sci. USA* **92**, 280–284 (1995).
- Schlötterer, C. & Tautz, D. Chromosomal homogeneity of *Drosophila* ribosomal DNA arrays suggests intrachromosomal exchanges drive concerted evolution. *Curr. Biol.* **4**, 777–783 (1994).

19. Arnheim, N., Treco, D., Taylor, B. & Eicher, E. M. Distribution of ribosomal gene length variants among mouse chromosomes. *Proc. Natl Acad. Sci. USA* **79**, 4677–4680 (1982).
20. Birky, C. W. Heterozygosity, heteromorphy, and phylogenetic trees in asexual eukaryotes. *Genetics* **144**, 427–437 (1996).
21. Hosny, M., Gianinazzi-Pearson, V. & Dulieu, H. Nuclear DNA content of 11 fungal species in Glomales. *Genome* **41**, 422–428 (1998).
22. Otto, S. P. & Whitton, J. Polyploid incidence and evolution. *Annu. Rev. Genet.* **34**, 401–437 (2000).
23. Mogie, M. & Ford, H. Sexual and asexual *Taraxacum* species. *Biol. J. Linn. Soc.* **35**, 155–168 (1988).
24. Kondrashov, A. S. The asexual ploidy cycle and the origin of sex. *Nature* **370**, 213–216 (1994).
25. Swofford, D. L. PAUP: phylogenetic analysis using parsimony (and other methods). Ver. 4 (Sinauer, Sunderland, MA, 1998).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank W. Pawlowski for help with the spore nuclei count; P. Bethke for advice on nuclei microdissection; D. Baker, T. Galagher, B. King, T. Lee and T. Szaro for technical assistance; J. A. Fortin for *G. intraradices*; D. Douds for the DC1 isolate of Ri T-DNA-transformed carrot roots developed by G. Bécard; and D. J. Read, T. Bruns, A. Burt and the members of the Taylor laboratory for comments on the manuscript. This work was supported by the Torrey Mesa Research Institute-Syngenta Biotechnology and the National Research Initiative Competitive Grants Program (NRI-CGP) of the US Department of Agriculture.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T.E.P. (tpawlows@nature.berkeley.edu). The sequences are deposited in GenBank under accession numbers AY330523–AY330580 (*G. etunicatum* PLS), AY330582–AY330597 (*G. etunicatum* ITS1–5.8S–ITS2 rDNA) and AY394030–AY394033 (*G. intraradices* ITS1 rDNA).

Multistability in the lactose utilization network of *Escherichia coli*

Ertugrul M. Ozbudak¹*, Mukund Thattai¹*, Han N. Lim¹, Boris I. Shraiman² & Alexander van Oudenaarden¹

¹Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

²Department of Physics and the BioMAPS Institute, Rutgers University, Piscataway, New Jersey 08854, USA

* These authors contributed equally to this work

Multistability, the capacity to achieve multiple internal states in response to a single set of external inputs, is the defining characteristic of a switch. Biological switches are essential for the determination of cell fate in multicellular organisms¹, the regulation of cell-cycle oscillations during mitosis^{2,3} and the maintenance of epigenetic traits in microbes⁴. The multistability of several natural^{1–6} and synthetic^{7–9} systems has been attributed to positive feedback loops in their regulatory networks¹⁰. However, feedback alone does not guarantee multistability. The phase diagram of a multistable system, a concise description of internal states as key parameters are varied, reveals the conditions required to produce a functional switch^{11,12}. Here we present the phase diagram of the bistable lactose utilization network of *Escherichia coli*¹³. We use this phase diagram, coupled with a mathematical model of the network, to quantitatively investigate processes such as sugar uptake and transcriptional regulation *in vivo*. We then show how the hysteretic response of the wild-type system can be converted to an ultrasensitive graded response^{14,15}. The phase diagram thus serves as a sensitive probe of molecular interactions and as a powerful tool for rational network design.

The bistability of the lactose utilization network has been under investigation since 1957 (refs 16, 17). The basic components of this network have been well characterized¹³, making it an ideal candidate

for global analysis. The *lac* operon comprises three genes required for the uptake and metabolism of lactose and related sugars (Fig. 1): *lacZ*, *lacY* and *lacA*. *lacZ* codes for β -galactosidase, an enzyme responsible for the conversion of lactose into allolactose and subsequent metabolic intermediates. *lacY* codes for the lactose permease (LacY), which facilitates the uptake of lactose and similar molecules, including thio-methylgalactoside (TMG), a non-metabolizable lactose analogue. *lacA* codes for an acetyltransferase, which is involved in sugar metabolism. The operon has two transcriptional regulators: a repressor (LacI) and an activator (the cyclic AMP receptor protein, CRP). Inducers, among them allolactose and TMG, bind to and inhibit repression by LacI, whereas cAMP binds to and triggers activation by CRP. The concentration of cAMP drops in response to the uptake of various carbon sources, including glucose and lactose¹⁸; glucose uptake also interferes with LacY activity, leading to exclusion of the inducer¹⁸. Together these effects mediate catabolite repression—the ability of glucose to inhibit *lac* expression. Crucially, cAMP levels are not affected by TMG uptake. Therefore, the extracellular concentrations of TMG and glucose can be used to independently regulate the activities of LacI and CRP, the two *cis*-regulatory inputs of the *lac* operon¹⁹. However, the response of the operon must be considered within the broader context of the network. The uptake of TMG induces the synthesis of LacY, which in turn promotes further TMG uptake; the resulting positive feedback loop creates the potential for bistability^{20,21}.

To probe the network's bistable response, we incorporated a single copy of the green fluorescent protein gene (*gfp*) under the control of the *lac* promoter into the chromosome of *E. coli* MG1655 (Fig. 1). We placed this reporter in the chromosome rather than on a multicopy plasmid to minimize the titration of LacI molecules by extraneous LacI-binding sites. The cells also contained a plasmid encoding a red fluorescent reporter (HcRed) under the control of the galactitol (*gat*) promoter. This promoter includes a CRP-binding site, as well as a binding site for the galactitol repressor,

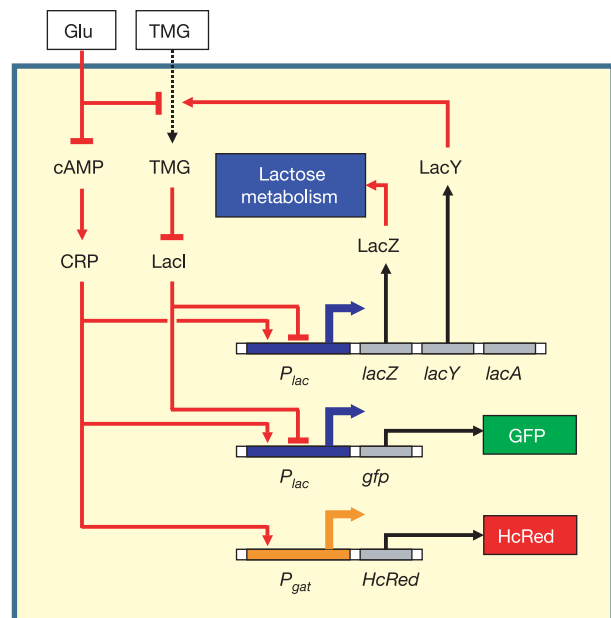


Figure 1 The lactose utilization network. Red lines represent regulatory interactions, with pointed ends for activation and blunt ends for inhibition; black arrows represent protein creation through transcription and translation, and dotted arrows represent uptake across the cell membrane. In our experiments we vary two external inputs, the extracellular concentrations of glucose and TMG, and measure the resulting levels of two fluorescent reporter proteins: GFP, expressed at the *lac* promoter, and HcRed, expressed at the *gat* promoter. LacY catalyses the uptake of TMG, which induces further expression of LacY, resulting in a positive feedback loop.

GatR. However, GatR is absent in *E. coli* MG1655 (ref. 22). Therefore, transcription at the *gat* promoter, measured by red fluorescence, is a direct measure of CRP-cAMP levels. In our experiments, we measure the response of single cells, initially in a given state of *lac* expression, to exposure to various combinations of glucose and TMG levels (Fig. 2a). It is crucial to use cells with well-defined initial states, either uninduced or fully induced, because the response of a bistable system is expected to depend on its history.

We find, in the absence of glucose, that the *lac* operon is uninduced at low TMG concentrations ($<3 \mu\text{M}$) and fully induced at high TMG concentrations ($>30 \mu\text{M}$) regardless of the cell's history. Between these switching thresholds, however, system response is hysteretic (history dependent): TMG levels must exceed $30 \mu\text{M}$ to turn on initially uninduced cells but must drop below $3 \mu\text{M}$ to turn off initially induced cells (Fig. 2b). As we approach the boundaries of this bistable region, stochastic mechanisms cause growing numbers of cells to switch from their initial states, resulting in a bimodal distribution of green fluorescence levels, with induced cells having over one hundred times the fluorescence levels of uninduced cells. This behaviour shows the importance of performing single-cell experiments, as a population-averaged measurement would have shown the mean fluorescence level to move smoothly between its low and high endpoints², obscuring the fact that individual cells never display intermediate fluorescence levels.

We find that the red fluorescence level is independent of cell history and of TMG concentration, showing that the observed history dependence of *lac* induction is not due to CRP. Red fluorescence levels do decrease in response to an increase in glucose concentration, ultimately dropping fivefold (Fig. 3a). There is a proportional drop in the green fluorescence levels of induced cells, reflecting the reduction in the levels of CRP-cAMP. The network continues to respond hysteretically in the presence of glucose, but higher levels of TMG are required to induce switching. By measuring system response at several glucose concentrations ranging from 0 to 1 mM, we are able to map out the complete range of glucose and TMG levels over which the system is bistable (Fig. 2c). This is the phase diagram of the wild-type lactose utilization network.

The switching boundaries of this phase diagram correspond to special conditions of network dynamics. By imposing these conditions within a mathematical model of the *lac* system, we are able to use the phase diagram as a quantitative probe of molecular processes in living single cells (see Supplementary Information).

The green fluorescence levels at each glucose and TMG concentration can be written as a function of three parameters: α , β and ρ . The maximal activity, α , is the *lac* expression level that would be obtained if every repressor molecule were inactive. The repression factor, ρ , gives the ratio of maximal to basal activities, the latter being the expression level that would be obtained if every repressor molecule were active. The transport rate, β , gives the TMG uptake rate per LacY molecule. We find that α is directly proportional to the red fluorescence level (Fig. 3b), demonstrating that the *lac* and *gat* promoters respond identically to CRP-cAMP. By contrast, α is essentially independent of TMG levels, suggesting that LacI and CRP bind independently to the *lac* operator site. We find the repression factor ρ to be 170 ± 30 regardless of the glucose and TMG levels (Fig. 3c). This confirms a strong prediction of our model—that ρ should be a function of the LacI concentration alone. Assuming an effective LacI concentration in the nanomolar range¹³, this value of ρ implies a dissociation constant between LacI and its major DNA-binding site of about 10^{-10} to 10^{-11} M, which is consistent with reported values²³.

The transport rate β is a product of two terms. The first term, β_T , which represents the TMG uptake rate per active LacY molecule, is purely a function of extracellular TMG levels (Fig. 3d). By fitting β_T to a hyperbola, we find that the half-saturation concentration for TMG uptake is $680 \pm 25 \mu\text{M}$, which agrees with previous measurements²⁴. The second term, β_G , which represents the fraction of LacY molecules that are active, is purely a function of extracellular glucose levels, reflecting the inducer-exclusion effect¹⁸. This allows us to separate catabolite repression into its constituent parts (Fig. 3e). We find that by lowering CRP-cAMP levels glucose reduces operon expression by 80%, whereas by inactivating LacY molecules it reduces TMG uptake by 35%. However, the network's positive feedback architecture amplifies these effects, resulting in the observed hundred-fold difference in *lac* expression levels between induced and uninduced cells. This type of global information would be extremely difficult to obtain using standard molecular-genetic techniques and *in vitro* biochemical assays. Our approach allows us to study the wild-type network in its entirety rather than isolated from other cellular systems or broken up into simpler subcomponents.

The phase diagram of the wild-type network (Fig. 2c) shows that *lac* induction always takes place hysteretically, with cells increasing their expression levels discontinuously as a switching threshold is

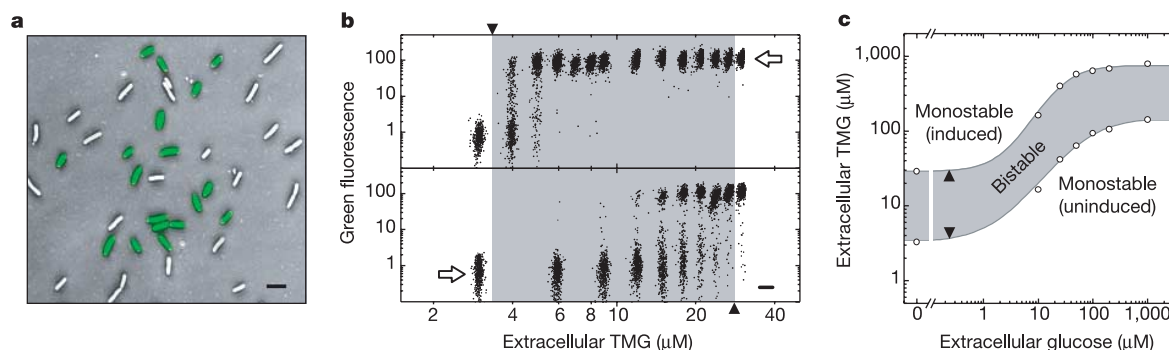


Figure 2 Hysteresis and bistability in single cells. **a**, Overlaid green fluorescence and inverted phase-contrast images of cells that are initially uninduced for *lac* expression, then grown for 20 h in $18 \mu\text{M}$ TMG. The cell population shows a bimodal distribution of *lac* expression levels, with induced cells having over one hundred times the green fluorescence of uninduced cells. Scale bar, $2 \mu\text{m}$. **b**, Behaviour of a series of cell populations, each initially uninduced (lower panel) or fully induced (upper panel) for *lac* expression, then grown in media containing various amounts of TMG. Scatter plots show log(green fluorescence) versus log(red fluorescence) for about 1,000 cells in each population. Each scatter plot is centred at a position that indicates the underlying TMG

concentration. The scale bar represents variation in red fluorescence by a factor of 10. White arrows indicate the initial states of the cell populations in each panel. The TMG concentration must increase above $30 \mu\text{M}$ to turn on initially uninduced cells (up arrow), whereas it must decrease below $3 \mu\text{M}$ to turn off initially induced cells (down arrow). The grey region shows the range of TMG concentrations over which the system is hysteretic. **c**, The phase diagram of the wild-type lactose utilization network. When glucose is added to the medium, the hysteretic region moves to higher levels of TMG. At each glucose level, the lower (down arrow) and upper (up arrow) switching thresholds show those concentrations of TMG at which less than 5% of the cells are in their initial states.

reached. However, our theoretical phase diagram (Fig. 4d) suggests that system response could also occur in a graded fashion, with the expression levels of individual cells moving continuously between low and high values. Such a response is predicted to occur when the degree of operon repression, ρ , is decreased below wild-type levels. Guided by this prediction, we sought to elicit a graded response from the *lac* system. We constructed two new strains, one containing a plasmid with an average copy number of 4, and the other containing a plasmid with an average copy number of 25. Each plasmid carried a single copy of the *lac* promoter. The introduction of extra LacI-binding sites had the expected titrating effect, reducing the effective LacI concentration and causing a drop in the operon-repression factor ρ from its wild-type value of 170, to 50 in the cells with 4 plasmids, and 5 in the cells with 25 plasmids. Graded behaviour is expected to occur below a repression factor of nine, and this is precisely what we observe: cells with a repression factor of 50 have a discontinuous hysteretic response similar to that of the original cells (Fig. 4b), whereas cells with a repression factor of 5 show a continuous graded response (Fig. 4c).

The ability of a single system to produce both binary and graded responses has previously been observed but not explained¹⁵. Here we present a general mechanism by which this may be achieved and experimentally validate this mechanism in the context of a natural biological system, the lactose utilization network. We show that the observed change from the wild-type binary response to the engineered graded response should be interpreted as a shift between different parts of a unified phase diagram. In this respect, the behaviour of the *lac* system closely resembles that of thermo-

dynamic systems¹¹: the discontinuous transition from low to high induction is analogous to a first-order phase transition such as evaporation in a liquid–gas system, with chemical noise instead of thermal noise driving stochastic transitions between these states^{8,25}. The shift from a hysteretic to a graded response is analogous to a second-order phase transition across a critical point. More complex transitions, such as those between different types of multistability²⁶ or between stable and oscillatory behaviour^{27,28}, can be similarly analysed. The response of any natural network may be regarded as a single realization in a vast but structured space of possible

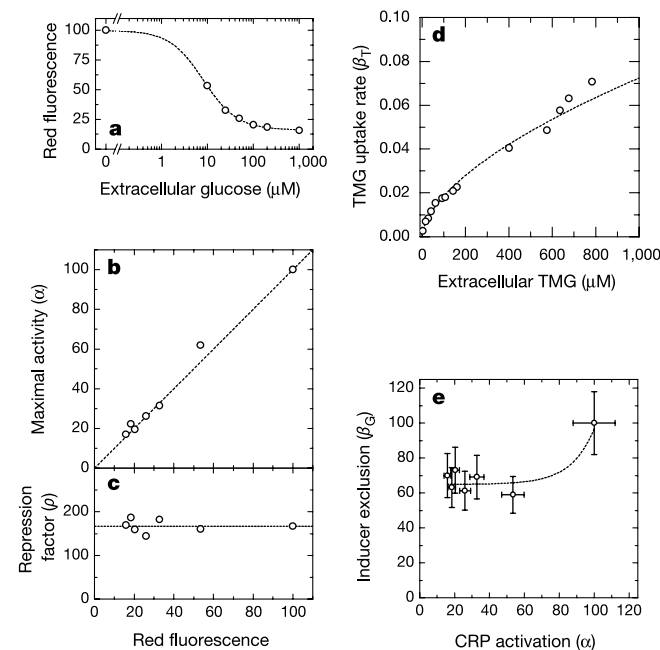


Figure 3 Single-cell *in vivo* measurement of network parameters. **a**, The mean red fluorescence level of each cell population is independent of its history but decreases with increasing glucose concentrations. **b–d**, Model parameters are extracted by fitting to measured fluorescence values at the switching thresholds. **b**, The maximal promoter activity, α , increases linearly with red fluorescence. **c**, The operon repression factor, ρ , is constant. **d**, The TMG uptake rate per active permease, β_T , increases with extracellular TMG levels. The dashed line shows a power-law fit with an exponent of 0.6. The data can also be fitted using a hyperbola, giving a half-saturation concentration of 680 μ M. **e**, The elements of catabolite repression. At each glucose concentration, we show the transport activity of LacY molecules (β_G , measuring inducer exclusion) versus the transcriptional activity of the *lac* operon (α , measuring CRP activation). We see that permease activity drops rapidly as glucose is added, falling to 65% of its maximal value. Operon activity drops more gradually, but falls to 20% of its maximal value. Error bars were determined by propagating the experimental error in measured fluorescence values.

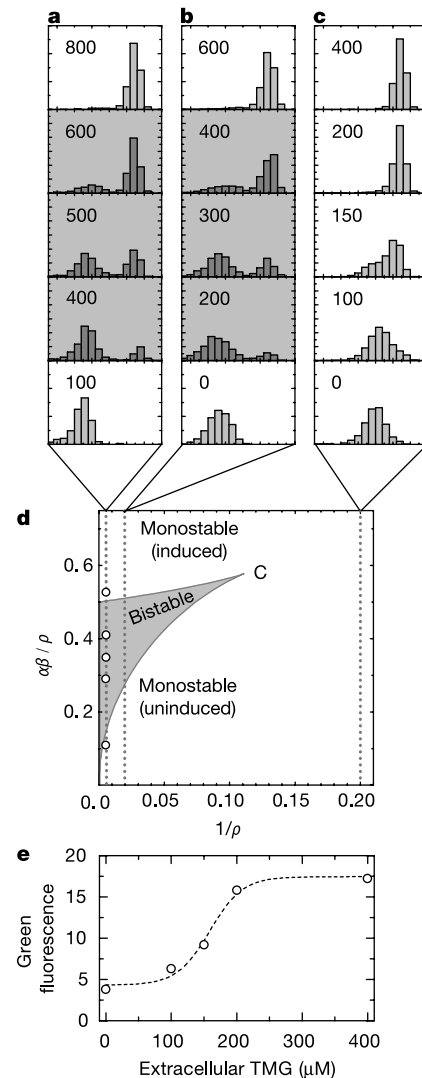


Figure 4 Hysteretic and graded responses. **a–c**, Histograms of log(green fluorescence) for cell populations that are initially uninduced, then grown in media containing 1 mM glucose and various levels of TMG (indicated in μ M on each panel). **a**, Response of the wild-type network. **b, c**, Response of the network after the introduction of extraneous LacI-binding sites on a 4-copy plasmid (**b**) and a 25-copy plasmid (**c**). The response of the wild-type and the 4-copy cells is hysteretic, whereas that of the 25-copy cells is graded. Bimodal histograms are shown in grey panels. **d**, Theoretical phase diagram of the lactose utilization network. It is possible to move from an uninduced to an induced state either hysteretically (grey region) or in a graded manner (white region). For the wild-type network we show the fitted parameters as white dots. For all three networks, we indicate the estimated operon-repression factors by vertical lines. The region of bistability grows smaller as the repression factor is decreased, eventually reaching a critical point (C) at a repression factor of 9. Graded behaviour is predicted to occur beyond this cusp. **e**, The mean green fluorescence level of each cell population in Fig. 4c is shown as a function of the TMG concentration. The response is highly sigmoidal (Hill coefficient $\approx$ 6) owing to positive feedback.

responses. By experimentally probing this space, we are able to gain quantitative insight into the architecture, dynamics and design constraints of biological systems. □

Methods

Bacterial strains and plasmids

The *gfp* gene under the control of the wild-type *lac* promoter, obtained from plasmid pGFPmut3.1 (Clontech), was inserted into the chromosome of *E. coli* MG1655 at the lambda insertion site using the λ -InCh technique²⁹ to produce the strain MUK21. The *gat* promoter was amplified from the *E. coli* MG1655 chromosome by polymerase chain reaction using primers flanking the 2,175,231–2,175,531 chromosomal region. The *HcRed* gene was obtained from pHcRed1–C1 (Clontech) and was placed under the control of the *gat* promoter into a plasmid with a ColE1 replication origin, which was transformed into MUK21 cells to obtain the strain ERT113. All measurements of wild-type network response were conducted in this strain. Two additional strains with *lac* operon repression factors at lower levels than in the wild-type were constructed by transforming MUK21 cells with multicopy plasmids, each incorporating a single copy of the *lac* promoter. The strain MUK21-pSC101* contains plasmids with a pSC101* replication origin (average copy number 4), and the strain MUK21-p15A contains plasmids with a p15A replication origin (average copy number 25)³⁰.

Growth conditions and media

Cells were grown at 37 °C in M9 minimal medium with succinate as the main carbon source, supplemented with varying amounts of glucose and TMG. Master cultures with cells induced for *lac* expression were prepared by overnight growth in 1 mM TMG, and master cultures with uninduced cells by overnight growth in the absence of TMG. During each experimental run, cells were transferred from these master cultures into media containing specified amounts of glucose and TMG. They were subsequently grown for 20 additional hours before they were harvested for measurement. The transfer volume was calculated to produce extremely low final cell densities ($OD_{600} \sim 0.001$), thereby preventing the depletion of glucose and TMG from the medium. Cells were concentrated by filtration and centrifugation, and the resulting pellet was resuspended in 2.5 μ l of the growth medium to prepare a microscope slide.

Data acquisition

Green and red fluorescence values of single cells were measured using a Nikon TE2000 microscope with automated stage and focus. For each experiment, images of about 1,000 cells on each slide were collected using a cooled back-thinned CCD camera (Micromax, Roper Scientific). These images were analysed using Metamorph (Universal Imaging) to obtain the average fluorescence of each cell above the fluorescence background.

Data analysis

For each glucose concentration, the fraction of cells in the induced state was determined as a function of TMG concentration, and the switching thresholds (defined as the TMG concentrations at which less than 5% of the cells are in their initial states) were obtained by interpolation. We estimated the green fluorescence values of the induced (high) and uninduced (low) subpopulations at each switching threshold by averaging over two neighbouring TMG concentrations. At each threshold, the high fluorescence was a linear function of the low fluorescence, with a small positive intercept comparable to the autofluorescence of *E. coli* MG1655 cells. We interpreted this intercept as the autofluorescence of the ERT113 strain. The repression factors for the MUK21-pSC101* and MUK21-p15A strains were estimated by taking the ratio of fully induced to uninduced fluorescence levels, assuming that these strains had the same autofluorescence background as ERT113.

Received 8 November; accepted 16 December 2003; doi:10.1038/nature02298.

1. Ferrell, J. E. Jr & Machleder, E. M. The biochemical basis of an all-or-none cell fate switch in *Xenopus* oocytes. *Science* **280**, 895–898 (1998).
2. Pomeroy, J. R., Sontag, E. D. & Ferrell, J. E. Jr Building a cell cycle oscillator: hysteresis and bistability in the activation of Cdc2. *Nature Cell Biol.* **5**, 346–351 (2003).
3. Sha, W. et al. Hysteresis drives cell-cycle transitions in *Xenopus laevis* egg extracts. *Proc. Natl Acad. Sci. USA* **100**, 975–980 (2003).
4. Hernday, A., Braaten, B. A. & Low, D. The mechanism by which DNA adenine methylase and PapI activate the pap epigenetic switch. *Mol. Cell* **12**, 947–957 (2003).
5. Blauwkamp, T. A. & Ninfa, A. J. Physiological role of the GlnK signal transduction protein of *Escherichia coli*: survival of nitrogen starvation. *Mol. Microbiol.* **46**, 203–214 (2002).
6. Siegle, D. A. & Hu, J. C. Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. *Proc. Natl Acad. Sci. USA* **94**, 8168–8172 (1997).
7. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).
8. Isaacs, F. J., Hasty, J., Cantor, C. R. & Collins, J. J. Prediction and measurement of an autoregulatory genetic module. *Proc. Natl Acad. Sci. USA* **100**, 7714–7719 (2003).
9. Becskei, A., Seraphin, B. & Serrano, L. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO J.* **20**, 2528–2535 (2001).
10. Ferrell, J. E. Jr Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Curr. Opin. Cell Biol.* **14**, 140–148 (2002).
11. Ma, S.-K. *Modern Theory of Critical Phenomena* (Perseus Books, Reading, Massachusetts, 1976).
12. Strogatz, S. H. *Nonlinear Dynamics and Chaos* (Perseus Books, Reading, Massachusetts, 1994).
13. Müller-Hill, B. *The Lac Operon: A Short History of a Genetic Paradigm* (Walter de Gruyter, Berlin, 1996).
14. Louis, M. & Becskei, A. Binary and graded responses in gene networks. *Science STKE* [online],

- 30 July 2002 (doi:10.1126/stke.2002.143.pe33).
15. Biggar, S. R. & Crabtree, G. R. Cell signaling can direct either binary or graded transcriptional responses. *EMBO J.* **20**, 3167–3176 (2001).
16. Novick, A. & Weiner, M. Enzyme induction as an all-or-none phenomenon. *Proc. Natl Acad. Sci. USA* **43**, 553–566 (1957).
17. Cohn, M. & Horibata, K. Inhibition by glucose of the induced synthesis of the β -galactoside-enzyme system of *Escherichia coli*: Analysis of maintenance. *J. Bacteriol.* **78**, 601–612 (1959).
18. Stulke, J. & Hillen, W. Carbon catabolite repression in bacteria. *Curr. Opin. Microbiol.* **2**, 195–201 (1999).
19. Setty, Y., Mayo, A. E., Surette, M. G. & Alon, U. Detailed map of a *cis*-regulatory input function. *Proc. Natl Acad. Sci. USA* **100**, 7702–7707 (2003).
20. Griffith, J. S. Mathematics of cellular control processes II: Positive feedback to one gene. *J. Theor. Biol.* **20**, 209–216 (1968).
21. Tyson, J. J. & Othmer, H. G. The dynamics of feedback control circuits in biochemical pathways. *Prog. Theor. Biol.* **5**, 1–62 (1978).
22. Nobelmann, B. & Lengeler, J. W. Molecular analysis of the *gat* genes from *Escherichia coli* and of their roles in galactitol transport and metabolism. *J. Bacteriol.* **178**, 6790–6795 (1996).
23. Oehler, S., Eismann, E. R., Kramer, H. & Müller-Hill, B. The three operators of the *lac* operon cooperate in repression. *EMBO J.* **9**, 973–979 (1990).
24. Chung, J. D. & Stephanopoulos, G. On physiological multiplicity and population heterogeneity of biological systems. *Chem. Eng. Sci.* **51**, 1509–1521 (1996).
25. Kepler, T. B. & Elston, T. C. Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys. J.* **81**, 3116–3136 (2001).
26. Thattai, M. & Shraiman, B. I. Metabolic switching in the sugar phosphotransferase system of *Escherichia coli*. *Biophys. J.* **85**, 744–754 (2003).
27. Atkinson, M. R., Savageau, M. A., Myers, J. T. & Ninfa, A. J. Development of genetic toggle circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*. *Cell* **113**, 597–607 (2003).
28. Smolen, P., Baxter, D. A. & Byrne, J. H. Frequency selectivity, multistability, and oscillations emerge from models of genetic regulatory systems. *Am. J. Physiol.* **43**, C531 (1998).
29. Boyd, D., Weiss, D. S., Chen, J. C. & Beckwith, J. Towards single-copy gene expression systems making gene cloning physiologically relevant: lambda InCh, a simple *Escherichia coli* plasmid-chromosome shuttle system. *J. Bacteriol.* **182**, 842–847 (2000).
30. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1–I2 regulatory elements. *Nucleic Acids Res.* **25**, 1203–1210 (1997).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank G. Jacobson and H. Kornberg, J. Paulsson, M. Savageau and A. Sengupta for discussions and suggestions; H. Bujard and R. Lutz for supplying the pZ vector system; and D. Boyd for help with the λ -InCh technique. We thank D. Raut for his assistance with the initial lactose measurements and the construction of plasmids and strains. We also thank A. Becskei and J. Pedraza for critically reviewing the manuscript. This work was supported by NIH and DARPA grants, and an NSF-CAREER grant.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.v.O. (avano@mit.edu).

Unique astrocyte ribbon in adult human brain contains neural stem cells but lacks chain migration

Nader Sanai^{1,2}, Anthony D. Tramontin^{1,2}, Alfredo Quiñones-Hinojosa¹, Nicholas M. Barbaro¹, Nalin Gupta¹, Sandeep Kunwar¹, Michael T. Lawton¹, Michael W. McDermott¹, Andrew T. Parsa¹, José Manuel-García Verdugo³, Mitchel S. Berger¹ & Arturo Alvarez-Buylla^{1,2}

¹Department of Neurological Surgery and Brain Tumor Research Center, and ²Developmental Stem Cell Biology Program, University of California San Francisco, San Francisco, California 94143, USA

³Instituto Cavanilles, University of Valencia, 46100, Spain

The subventricular zone (SVZ) is a principal source of adult neural stem cells in the rodent brain, generating thousands of olfactory bulb neurons every day¹. If the adult human brain contains a comparable germinal region, this could have considerable implications for future neuroregenerative therapy. Stem cells have been isolated from the human brain^{2–7}, but the identity, organization and function of adult neural stem cells in the

human SVZ are unknown. Here we describe a ribbon of SVZ astrocytes lining the lateral ventricles of the adult human brain that proliferate *in vivo* and behave as multipotent progenitor cells *in vitro*. This astrocytic ribbon has not been observed in other vertebrates studied. Unexpectedly, we find no evidence of chains of migrating neuroblasts in the SVZ or in the pathway to the olfactory bulb. Our work identifies SVZ astrocytes as neural stem cells in a niche of unique organization in the adult human brain.

Young neurons born in the rodent and primate SVZ migrate in chains through the rostral migratory stream (RMS) to replace interneurons of the olfactory bulb^{8–10}. Data from rodents indicate that SVZ neural stem cells are astrocytes^{11–14}. In the presence of exogenous mitogens, SVZ-derived neural stem cells self-renew and can differentiate into astrocytes, oligodendrocytes and neurons *in vitro*^{15–17}. Explants of intraoperative and postmortem adult human brain also contain progenitors that can give rise to new neurons and glia^{2,4,6,7,18,19}, but the organization and function of the human germinal zones that contain these progenitors remain poorly understood.

We collected adult human SVZ specimens from 65 neurosurgical resections and 45 autopsied brains (Fig. 1a; specimens are listed in Supplementary Tables 1 and 2). With a starting concentration of

50,000 cells per ml (refs 15, 17), intraoperative samples from the anterior horn, body, atrium, occipital horn and temporal horn of the lateral ventricular walls generated on average 62.57 ± 7.46 (mean $\pm$ s.d.) neurospheres per ml in the presence of epidermal growth factor (EGF) and fibroblast growth factor (FGF). On average, 33.11 ± 3.81 secondary neurospheres were generated from 100 primary neurospheres.

Human SVZ neurospheres differentiated into glia fibrillary acidic protein (GFAP)-positive astrocytes, O4-positive oligodendrocytes and TuJ1-positive neurons. Primary neurospheres ($n = 35$ independent samples) yielded $6.5 \pm 3.4\%$ neurons, $5.8 \pm 1.5\%$ oligodendrocytes and $87.8 \pm 3.1\%$ astrocytes. In parallel, specimens that excluded the SVZ, including parietal cortex ($n = 8$), temporal cortex ($n = 10$), striatum ($n = 14$), medial (septal) wall of the lateral ventricle ($n = 3$), third ventricle ($n = 2$) and fourth ventricle ($n = 1$), did not generate neurospheres. Our results indicate that EGF- and FGF-responsive precursors exist not only in the temporal lobe^{4,7}, but in samples from different regions of the lateral walls of the human lateral ventricles.

Sections of adult human SVZ specimens ($n = 68$) showed a band of cells separated from the ependyma (Fig. 1b). Most cells in this band were astrocytes, showing stellate morphology and expressing GFAP (Fig. 1c) and vimentin (data not shown). This astrocyte ribbon was present in intraoperative and autopsied specimens from the lateral ventricular walls of male and female individuals aged 19–68 yr. It was not observed, however, along the third or fourth ventricles ($n = 3$), nor was it present in samples of the medial (septal) wall of the lateral ventricles ($n = 6$). This astroglia band has not been described in the SVZ of primates^{8,10}, rodents²⁰, dogs²¹, cows²² or sheep (unpublished observations), and represents a departure from classical SVZ cytoarchitecture²⁰.

Under the electron microscope, cells in the band had ultrastructural characteristics of astrocytes (Figs 1d and 2a). Astrocytes in the SVZ ribbon were separated from the ependyma by a gap filled with GFAP-positive processes (Supplementary Fig. 1). These processes contained intermediate filaments and gap junctions (Fig. 2b), confirming their astrocytic origin. No chains of migratory neuroblasts were seen in the ribbon or gap, although occasionally single, elongated, TuJ1-positive cells were observed (Supplementary Fig. 3b). A few astrocyte somata were also found in the gap region (Fig. 2c). Notably, some of these astrocytes had a process extending into ependyma, towards the ventricular lumen (Fig. 2d).

Some cells in the astrocyte ribbon coexpressed GFAP and the cell-division marker Ki-67, suggesting that a subpopulation of them were mitotically active *in vivo* (Fig. 2e). To confirm this observation, organotypic slices of human SVZ ($n = 3$) were incubated in medium containing 5-bromodeoxyuridine (BrdU). After 36 h, an antibody against BrdU was found to colocalize with GFAP-positive cells in the SVZ ribbon (Supplementary Fig. 2). In a study of individuals treated with BrdU, proliferative SVZ cells were found in the same position²³. Taken together, these findings indicate that cells in this band of SVZ astrocytes divide *in vivo*. From Ki-67 labelling ($n = 10$ total specimens), we estimated that $0.77 \pm 0.29\%$ of astrocytes in the SVZ ribbon were in division. Cultures of microdissected adult human SVZ yielded monolayers of astrocytes. Notably, in comparison to cortical astrocyte cultures, SVZ astrocytes grew at a considerably higher rate (Fig. 3a).

We next examined whether adult human SVZ astrocytes could produce multipotent, self-renewing neurospheres. Cultures of SVZ astrocytes ($n = 25$ specimens) were dissociated and incubated in serum-free medium containing EGF and basic FGF (bFGF). Seeding at an initial concentration of 50,000 astrocytes per ml, we generated 109.29 ± 8.67 neurospheres per ml.

We then tested whether single human SVZ astrocytes could produce multipotent, self-renewing neurospheres. SVZ astrocyte cultures were established from the lateral walls of the anterior horn ($n = 3$), body ($n = 2$) and temporal horn ($n = 3$). Cells were

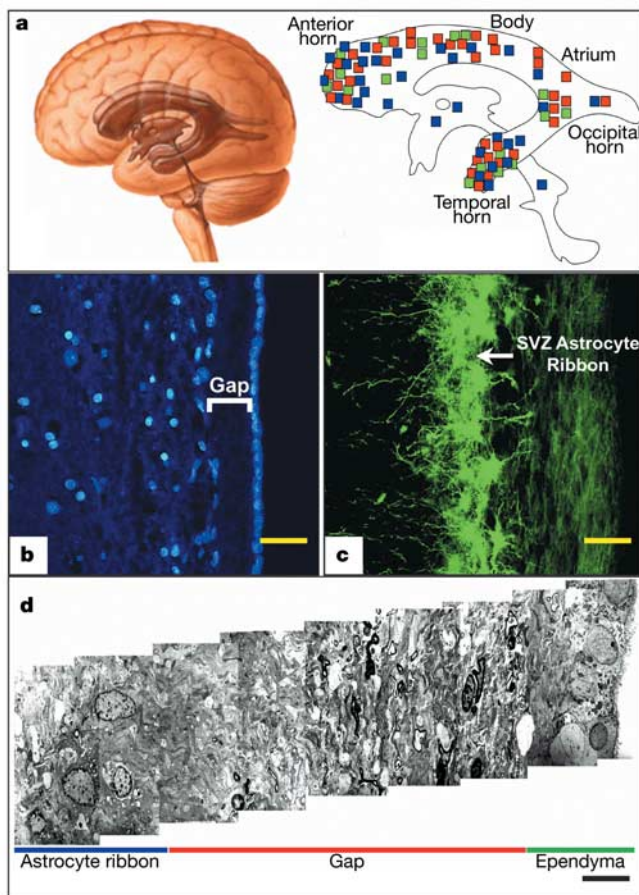


Figure 1 Dense ribbon of SVZ astrocytes in the adult human brain. **a**, SVZ specimen locations along lateral wall of the human ventricles. Intraoperative tissue was used for histology (red) and culture (green) experiments. Autopsied tissue (blue) was used only for histology. **b**, Coronal section (6 μ m) stained with the nuclear marker DAPI reveals a region of high cellularity that is separated from the ependyma by a gap. **c**, Vibratome section showing GFAP expression of SVZ astrocyte ribbon and GFAP-positive fibre bundles filling the subependymal gap. **d**, Panoramic electron micrograph of a postmortem adult human SVZ, showing the ependymal lining, gap region and astrocyte ribbon. Scale bars, 40 μ m (**b**, **c**); 10 μ m (**d**).

incubated for 3 d with an adenovirus carrying the gene for green fluorescent protein (GFP) controlled by the GFAP promoter (GFAPp)¹¹. Single green astrocytes were individually micropipetted into separate wells of conditioned neurosphere medium. Of 121 anterior horn SVZ astrocytes plated, 5 (4.1%) generated clonal neurospheres. Similarly, of 113 ventricle body SVZ astrocytes and 137 temporal horn SVZ astrocytes, 4 (3.5%) and 4 (2.9%) clonal neurospheres were produced, respectively. Dissociation of the neurospheres derived from single astrocytes resulted in the production of secondary neurospheres. On the differentiation of primary and secondary neurospheres, we identified astrocytes, oligodendrocytes and neurons in each colony (Fig. 3b). Single cortical astrocytes (139 tested) and single astrocytes derived from tissue adjacent to the SVZ (201 tested) did not generate neurospheres. These experiments indicate that human SVZ neural stem cells correspond to astrocytes.

The high concentrations of exogenous mitogens required for neurosphere formation may amplify the true capabilities of a cell. We therefore tested whether primary SVZ astrocytes generate neurons without EGF, bFGF or neurosphere formation. In the absence of exogenous EGF and FGF, a substrate of astrocytes can support neurogenesis for adult neural stem cells^{24,25}. Cultured SVZ astrocytes from the anterior horn ($n = 8$ specimens) were incubated with the GFAPp–GFP adenovirus and the nuclear marker 4',6-diamidino-2-phenylindole dihydrochloride (DAPI). Single green astrocytes were placed onto human cortical astrocyte monolayers in serum-free medium (Fig. 3c, inset). At 5–14 d *in vitro*, 5 out of 64 (7.8%) SVZ astrocytes generated colonies with bipolar and/or multipolar cells expressing DAPI and TuJ1 (Fig. 3c). By contrast, single cortical astrocytes ($n = 80$) placed under identical conditions yielded only astrocytes. These observations provide direct evidence that single adult human SVZ astrocytes, in the absence of exogenous growth factors, are capable of neurogenesis.

New neurons born in the SVZ migrate in chains along the RMS.

This migration is well documented in many adult mammalian species^{26,27}, including primates^{8,10}, but not in humans. The large pool of SVZ neuronal progenitors that we describe suggested that a similar migration might occur in adult humans. However, we found no evidence of neuronal chain migration in the SVZ. Serial coronal and sagittal sections ($n = 9$ total specimens; Fig. 4a) did reveal a structure beginning at the ventral floor of the anterior horn and coursing ventro-rostrally to the olfactory trigone—the origin of the human olfactory peduncle (Fig. 4b). This SVZ–olfactory trigone connection contained displaced ependymal cavities (Fig. 4b, inset), hallmarks of descent from an embryonic olfactory ventricle that precedes the adult RMS. Nissl staining and double labelling for TuJ1 and polysialylated neural cell adhesion molecule (PSA-NCAM)^{8,10} did not show evidence of chains of young migratory neurons in this region. Cells in this human SVZ–olfactory trigone connection were rich in GFAP (Fig. 4c), suggesting that this region contains many astrocytes.

Regardless of their route out of the SVZ, chains of newly generated neurons, if present, must traverse the human olfactory peduncle to reach the bulbs. Under the light microscope, the olfactory peduncle was composed of a network of myelinated axons, corpus amylaceum²⁸, astrocytes and GFAP-positive processes (Fig. 4d, e, and Supplementary Fig. 4c), but devoid of TuJ1-positive or PSA-NCAM-positive chain-migrating neurons (Fig. 4f and Supplementary Fig. 4b). Rarely, an elongated cell expressing TuJ1 was observed, but these cells were not organized into chains. Expression of TuJ1 in this and other human brain regions (Supplementary Fig. 3a) served as a positive control, indicating that any lack of labelling was not due to antibody incompatibility with human material. We confirmed these results by electron microscopy of serial semithin sections of intraoperative and postmortem olfactory peduncle ($n = 12$ total specimens; Fig. 4d and Supplementary Fig. 4a). Again, we found no evidence of chains of migrating neurons in the olfactory peduncle or a structure similar

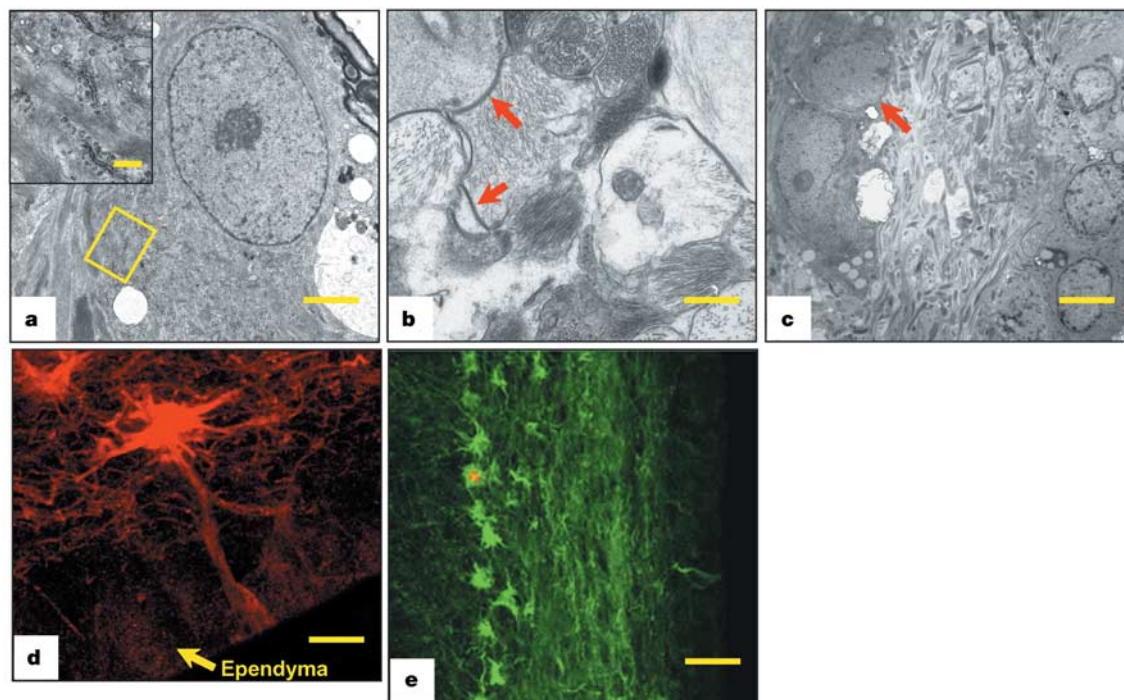


Figure 2 Characteristics of adult human SVZ astrocytes. **a**, Electron micrograph of the SVZ ribbon identifies an astrocyte nucleus with a cell body containing copious intermediate filaments (inset) and rough endoplasmic reticulum. **b**, SVZ gap region contains many astrocytic processes and a network of gap junctions. **c**, Nuclei (red arrows) of two displaced astrocytes in the gap region. Notice how these cells are surrounded by

other astrocytic processes in the gap. Nearby ependymal lining is seen on the right. **d**, Displaced SVZ astrocyte soma extends a process towards the lateral ventricle. **e**, Ki-67 labelling (red) colocalizing to a proliferating GFAP-positive astrocyte (green) in the SVZ ribbon. Scale bars, 1.5 μm (**a**); 0.1 μm (**a**, inset); 1 μm (**b**); 3 μm (**c**); 7 μm (**d**); 40 μm (**e**).

to the RMS of other mammals. These findings raise the unexpected possibility that migration from the SVZ to the olfactory bulb does not take place in adult humans or, if it does, precursors migrate as individual cells.

We have described a subventricular organization in the adult human brain that differs considerably from that of other vertebrates that have been studied. This germinal region contains cells with astrocytic characteristics that can function as neural stem cells and

features a previously uncharacterized periventricular astrocyte ribbon. Although the precise location of the stem cells cannot be established, we know that astrocytes from this region are clonal precursors of self-renewing, multipotent neurospheres and can give rise to neurons in the absence of exogenous growth factors. Considering the size of the human lateral ventricular system, our work suggests that a substantial number of neural stem cells exist in the adult brain throughout life.

Despite such robust germinal capacity, there is no evidence of cells migrating in chains along the SVZ or olfactory peduncle to the bulb. This may be explained by the long distance that separates the olfactory bulbs from the cerebrum in humans and/or our relatively micro-osmotic capabilities. Our findings raise the question of the fate of cells born in the human SVZ. The identification of a large, anatomically discrete population of proliferating, multipotent human astrocytes may also lead to a better understanding of the role that human neural stem cells have in tumorigenesis, demyelination and neurodegeneration. This work highlights the importance of studying stem cells in the human brain. □

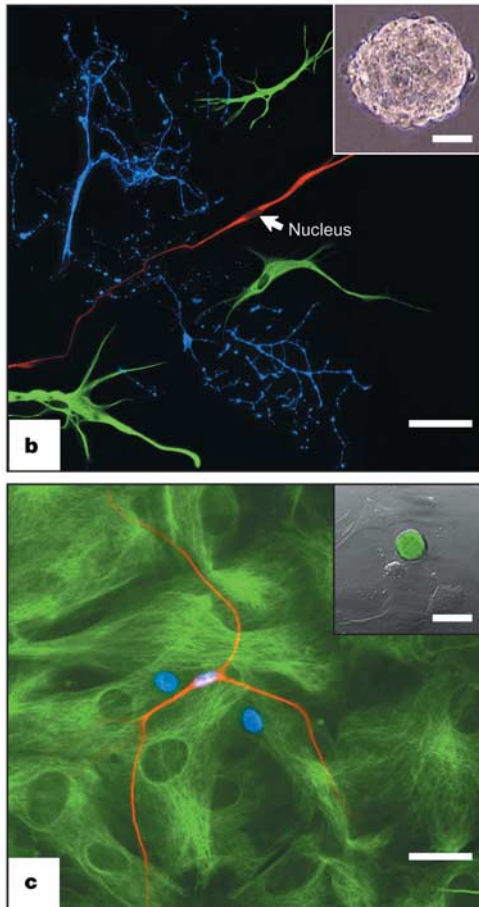
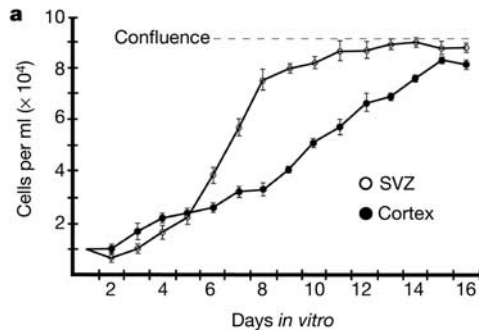


Figure 3 Human SVZ astrocytes as neural progenitors. **a**, Proliferation rates for SVZ and cortical astrocytes were significantly different ($P < 0.05$). **b**, Single SVZ astrocytes generated clonal neurospheres (inset) that differentiate into TuJ1-positive neurons (red), GFAP-positive astrocytes (green) and O4-positive oligodendrocytes (blue). **c**, Single SVZ astrocytes transfected with a GFAP-GFP adenovirus fluoresced green and were placed on a cortical astrocyte monolayer; inset shows superimposed differential interference contrast and green fluorescence images. SVZ astrocyte progeny were traced with DAPI (blue). The resultant DAPI-stained colony grows on GFAP-positive cortical astrocytes (green) and yields a DAPI-positive cell colocalizing with the young neuronal marker TuJ1 (red), shown here at 10 d *in vitro*. Scale bars, 80 μ m (**b**); 100 μ m (**b** inset, **c**); 60 μ m (**c** inset).

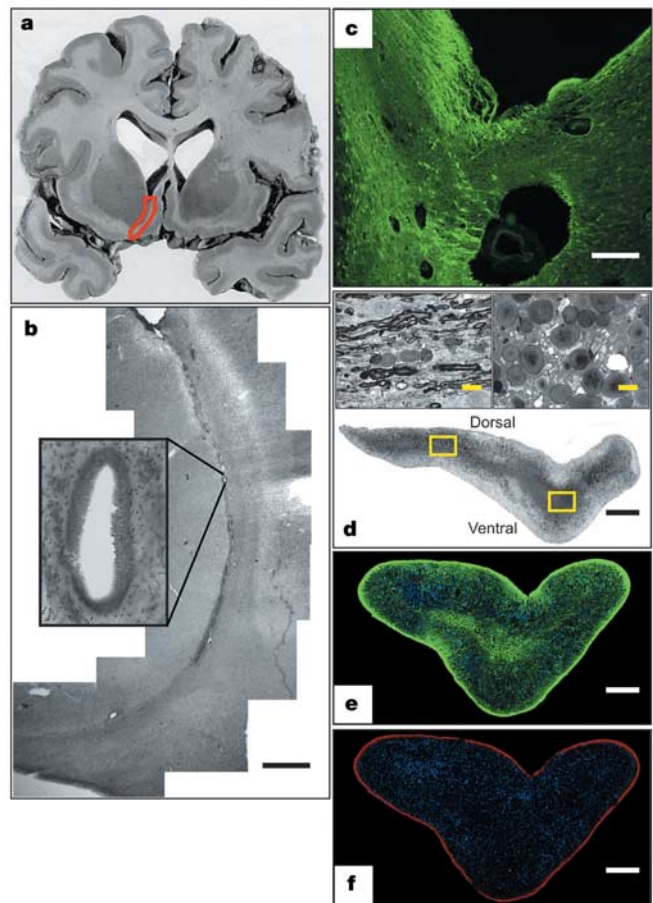


Figure 4 The human SVZ-olfactory trigone connection: no evidence for neuronal chain migration to the olfactory bulbs. **a**, Pathology specimen of the human brain region corresponding to the location of the mammalian RMS (red). **b**, Nissl staining of the human SVZ-olfactory trigone connection. Ependymal cavities (inset) suggest its embryonic ventricular origins. **c**, GFAP-positive fibres converge from the ventral floor of the anterior horn towards the olfactory trigone. **d**, Transverse section of the adult human olfactory peduncle. The left-hand box is magnified in the left inset, showing myelinated axons and oligodendrocytes; the right-hand box is magnified in the right inset, showing corpus amylaceum bodies found in the core of the olfactory peduncle. **e**, TuJ1 (red) and DAPI (blue) labelling of transverse olfactory peduncle, showing no evidence of chains of young neurons migrating to the olfactory bulb. **f**, GFAP (green) and DAPI (blue) labelling of transverse olfactory peduncle, showing abundant astrocytic processes, but no RMS. Scale bars, 5 mm (**b**); 500 μ m (**c**, **d**); 25 μ m (**d** insets); 400 μ m (**e**, **f**).

Methods

Intraoperative and pathology specimens

Neurosurgical excisions of normal SVZ occurred as part of the planned margin of resection surrounding a periventricular lesion (Supplementary Table 1). We recorded the anatomical origin of each intraoperative specimen with stealth neuronavigation²⁹. Intraoperative specimens were histologically normal with no evidence of dysplasia, and assessments were independently confirmed by a neuropathologist. For pathological specimens, autopsied brains were cut coronally at the temporal horn tip and immersed in 3% paraformaldehyde (PFA) or formalin for 1–2 weeks, and then portions of their lateral ventricular walls were excised (Supplementary Table 2). Autopsy specimens were obtained within 12 h of death from individuals who lacked clinical or postmortem evidence of brain pathology. All causes of death were non-neurological.

Expression of TuJ1 and PSA-NCAM was noted in different areas of brain parenchyma, indicating that the antibodies used recognized human antigen and that the human brain's antigenicity was preserved (Supplementary Fig. 3a, b). The *in vitro* and *in vivo* characteristics of human SVZ were observed in samples from individuals with different lesions and of different ages and sexes, and in intraoperative and pathology specimens alike, making it extremely unlikely that our results were due to specific diseases. Collection of all specimens was done in accordance with the University of California San Francisco Committee on Human Research.

Immunohistochemistry

All fixed specimens were rinsed in 0.1 M PBS and then cut on a vibratome (50 μ m), cryostat (10 μ m) or microtome (6 μ m). Tissue sections were incubated overnight at 4 °C in blocking solution (PBS, 0.25% Triton X-100 and 10% normal donkey serum). Sections were then incubated for 48 h at 4 °C in primary antibody diluted in blocking solution. Primary antibodies dilutions were as follows: anti-GFAP (Chemicon), 1:500; anti-TuJ1 (Covance), 1:500; anti-Ki-67 (clone MIB-1, Dako), 1:50; anti-vimentin (Developmental Studies Hybridoma Bank), 1:1; anti-BrdU (Oxford Biotech), 1:10; and anti-PSA-NCAM (a gift from G. Rougon), 1:3,000. The antibody against O4 (Chemicon) was diluted to 1:100 in blocking solution without Triton X-100. After primary antibody incubation, sections were incubated for 18 h at 4 °C in secondary antibody (Jackson Laboratories) diluted 1:200. We used DAPI (1:5,000) for counterstaining, and Aqua Polymount (Polysciences) for mounting. In all experiments, omitting the primary antibody resulted in no labelling. Positive controls of rodent SVZ and human cortical specimens were analysed in parallel with experimental specimens.

Cell culture

Complete medium was a 0.2- μ m filter-sterilized mixture of 44 ml of DMEM medium (Gibco), 5 ml of fetal calf serum, 0.5 ml of antibiotic or antimycotic and 0.5 ml of L-glutamine. After resection, intraoperative specimens were minced into 50–300- μ m pieces, incubated in 0.25% trypsin and 0.02% versene for 10 min, triturated and plated onto uncoated T-25 flasks with complete medium. At confluence, flasks were shaken for 16 h at 300 r.p.m. to remove neurons, oligodendrocytes and microglia. This method³⁰ has been observed to yield cell cultures that are >98% GFAP positive²⁴.

Organotypic explant

Intraoperative human SVZ specimens were dissected in ice-cold complete medium under a microscope. Each 200- μ m specimen was floated on a Millipore filter in a 35-mm dish with complete medium plus BrdU (10 ng μ l⁻¹). After 36 h, specimens were immersed in 3% PFA for 24 h at 4 °C and then sectioned by cryostat. Immunohistochemistry was done as described above.

Neurosphere assay

Intraoperative specimens were dissociated as above and incubated in non-adherent neurosphere conditions as described¹⁷. Each 1-cm² well contained 1 ml of neurosphere medium¹⁷ and a cell concentration of 50 cells per μ l. Neurospheres were passaged at least three times without a significant difference in plating efficiency. Conditioned neurosphere medium was used only in clonal neurosphere assays, for which complete neurosphere medium was exposed to mature neurospheres for 24 h before the cells were filtered out. The resultant medium was combined in a 50:50 mixture with fresh complete neurosphere medium to generate conditioned neurosphere medium.

Cell division and astrocyte growth curve

The quantification of human astrocyte proliferation is described in the Supplementary Information.

Clonal monolayer assay

Single astrocytes, identified after 3 d of incubation with a GFAPp-GFP adenovirus¹¹, were incubated in DAPI (1:5,000) for 30 min at 37 °C and placed on a cortical astrocyte monolayer in 16-well glass slides with Neurobasal/B27 medium (Life Technologies). On the basis of immunohistochemistry, the specificity of the GFAPp-GFP adenovirus was >98% in human astrocyte cultures. Neurobasal/B27 medium was a mixture of 44 ml of DMEM-F12 medium, 5 ml of B27 supplement, 0.5 ml of L-glutamine and 0.5 ml of antibiotic or antimycotic. At 5–14 d *in vitro*, monolayers were fixed in PFA and processed with immunohistochemistry.

Electron microscopy

Specimens fixed in 2% glutaraldehyde and 2% paraformaldehyde were cut into 200- μ m sections on a vibratome. Sections were post-fixed in 2% osmium, rinsed, dehydrated and embedded in Araldite (Durcupan, Fluka). To study SVZ architecture, we cut serial 1- μ m

semithin sections and stained them with 1% toluidine blue. To identify individual cell types, ultrathin (0.05- μ m) sections were cut with a diamond knife, stained with lead citrate and examined under a Jeol 100CX electron microscope.

Received 24 June; accepted 19 December 2003; doi:10.1038/nature02301.

1. Alvarez-Buylla, A., Garcia-Verdugo, J. M. & Tramontin, A. D. A unified hypothesis on the lineage of neural stem cells. *Nature Rev. Neurosci.* **2**, 287–293 (2001).
2. Pincus, D. W. *et al.* *In vitro* neurogenesis by adult human epileptic temporal neocortex. *Clin. Neurosurg.* **44**, 17–25 (1997).
3. Roy, N. S. *et al.* Promoter-targeted selection and isolation of neural progenitor cells from the adult human ventricular zone. *J. Neurosci. Res.* **59**, 321–331 (2000).
4. Johansson, C. B., Svensson, M., Wallstedt, L., Janson, A. M. & Frisen, J. Neural stem cells in the adult human brain. *Exp. Cell Res.* **253**, 733–736 (1999).
5. Pagano, S. F. *et al.* Isolation and characterization of neural stem cells from the adult human olfactory bulb. *Stem Cells* **18**, 295–300 (2000).
6. Nunes, M. C. *et al.* Identification and isolation of multipotential neural progenitor cells from the subcortical white matter of the adult human brain. *Nature Med.* **9**, 439–447 (2003).
7. Kukekov, V. G. *et al.* Multipotent stem/progenitor cells with similar properties arise from two neurogenic regions of adult human brain. *Exp. Neurol.* **156**, 333–344 (1999).
8. Kornack, D. R. & Rakic, P. The generation, migration, and differentiation of olfactory neurons in the adult primate brain. *Proc. Natl Acad. Sci. USA* **98**, 4752–4757 (2001).
9. Alvarez-Buylla, A. & Garcia-Verdugo, J. M. Neurogenesis in adult subventricular zone. *J. Neurosci.* **22**, 629–634 (2002).
10. Pencea, V., Bingaman, K. D., Freedman, L. J. & Luskin, M. B. Neurogenesis in the subventricular zone and rostral migratory stream of the neonatal and adult primate forebrain. *Exp. Neurol.* **172**, 1–16 (2001).
11. Doetsch, F., Caille, L., Lim, D. A., Garcia-Verdugo, J. M. & Alvarez-Buylla, A. Subventricular zone astrocytes are neural stem cells in the adult mammalian brain. *Cell* **97**, 703–716 (1999).
12. Laywell, E. D., Rakic, P., Kukekov, V. G., Holland, E. C. & Steindler, D. A. Identification of a multipotent astrocytic stem cell in the immature and adult mouse brain. *Proc. Natl Acad. Sci. USA* **97**, 13883–13888 (2000).
13. Skogh, C. *et al.* Generation of regionally specified neurons in expanded glial cultures derived from the mouse and human lateral ganglionic eminence. *Mol. Cell. Neurosci.* **17**, 811–820 (2001).
14. Imura, T., Kornblum, H. I. & Sofroniew, M. V. The predominant neural stem cell isolated from postnatal and adult forebrain but not early embryonic forebrain expresses GFAP. *J. Neurosci.* **23**, 2824–2832 (2003).
15. Reynolds, B. A. & Weiss, S. Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system. *Science* **255**, 1707–1710 (1992).
16. Chiasso, B. J., Tropepe, V., Morshead, C. M. & van der Kooy, D. Adult mammalian forebrain ependymal and subependymal cells demonstrate proliferative potential, but only subependymal cells have neural stem cell characteristics. *J. Neurosci.* **19**, 4462–4471 (1999).
17. Gritti, A. *et al.* Multipotent neural stem cells reside into the rostral extension and olfactory bulb of adult rodents. *J. Neurosci.* **22**, 437–445 (2002).
18. Kirschenbaum, B. *et al.* *In vitro* neuronal production and differentiation by precursor cells derived from the adult human forebrain. *Cereb. Cortex* **4**, 576–589 (1994).
19. Palmer, T. D. *et al.* Cell culture. Progenitor cells from human brain after death. *Nature* **411**, 42–43 (2001).
20. Doetsch, F., Garcia-Verdugo, J. M. & Alvarez-Buylla, A. Cellular composition and three-dimensional organization of the subventricular germinal zone in the adult mammalian brain. *J. Neurosci.* **17**, 5046–5061 (1997).
21. Blakemore, W. F. & Jolly, R. D. The subependymal plate and associated ependyma in the dog. An ultrastructural study. *J. Neurocytol.* **1**, 69–84 (1972).
22. Perez-Martin, M. *et al.* Ependymal explants from the lateral ventricle of the adult bovine brain: a model system for morphological and functional studies of the ependyma. *Cell Tissue Res.* **300**, 11–19 (2000).
23. Eriksson, P. S. *et al.* Neurogenesis in the adult human hippocampus. *Nature Med.* **4**, 1313–1317 (1998).
24. Lim, D. A. & Alvarez-Buylla, A. Interaction between astrocytes and adult subventricular zone precursors stimulates neurogenesis. *Proc. Natl Acad. Sci. USA* **96**, 7526–7531 (1999).
25. Song, H., Stevens, C. F. & Gage, F. H. Astroglia induce neurogenesis from adult neural stem cells. *Nature* **417**, 39–44 (2002).
26. Lois, C., Garcia-Verdugo, J. M. & Alvarez-Buylla, A. Chain migration of neuronal precursors. *Science* **271**, 978–981 (1996).
27. Bonfanti, L., Peretto, P., Merighi, A. & Fasolo, A. Newly-generated cells from the rostral migratory stream in the accessory olfactory bulb of the adult rat. *Neuroscience* **81**, 489–502 (1997).
28. Leel-Ossy, L. & Gati, I. Corpus amylaceum (polyglucosan body) in the peripheral olfactory system. *Pathol. Oncol. Res.* **4**, 212–216 (1998).
29. Smith, K. R., Frank, K. J. & Bucholz, R. D. The NeuroStation—a highly accurate, minimally invasive solution to frameless stereotactic neurosurgery. *Comput. Med. Imaging Graph.* **18**, 247–256 (1994).
30. McCarthy, K. D. & de Vellis, J. Preparation of separate astroglial and oligodendroglial cell cultures from rat cerebral tissue. *J. Cell Biol.* **85**, 890–902 (1980).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank G. E. Vates and J. W. Chiong for comments on the manuscript; A. Bollen, D. Gaskin, T. Tihan and S. R. VandenBerg for assistance with pathology material; A. Leong for stealth neuronavigation imagery; and Z. Mirzadeh for imaging. This work was supported by a gift from Frances and John Bowes, by the William Siebrandt Glioblastoma Fund and grants from the NIH. A.A.-B. is the Heather and Melanie Muss Professor. N.S. is supported by the HHMI and an American Brain Tumor Association David B. Anderson Fellowship. A.D.T. is supported by the Damon Runyon Cancer Research Foundation.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to N.S. (nsanai@itsa.ucsf.edu) or A.A.-B. (abuylla@itsa.ucsf.edu).

Direct interaction of geminin and Six3 in eye development

Filippo Del Bene, Kristin Tessmar-Raible & Joachim Wittbrodt

Developmental Biology Programme, EMBL, Meyerhofstrasse 1, 69012 Heidelberg, Germany

Organogenesis in vertebrates requires the tight control of cell proliferation and differentiation. The homeobox-containing transcription factor Six3 plays a pivotal role^{1,2} in the proliferation of retinal precursor cells. In a yeast two-hybrid screen, we identified the DNA replication-inhibitor geminin as a partner of Six3. Geminin inhibits cell-cycle progression³ by sequestering Cdt1 (refs 4, 5), the key component for the assembly of the pre-replication complex⁶. Here, we show that Six3 efficiently competes with Cdt1 directly to bind to geminin, which reveals how Six3 can promote cell proliferation without transcription. In common with Six3 inactivation^{2,7}, overexpression of the geminin gene (*Gem*; also known as *Gmn*) in medaka (*Oryzias latipes*) induces specific forebrain and eye defects that are rescued by Six3. Conversely, loss of *Gem* (in common with gain of Six3 (ref. 1)) promotes retinal precursor-cell proliferation and results in expanded optic vesicles, markedly potentiating Six3 gain-of-function phenotypes. Our data indicate that the transcription factor Six3 and the replication-initiation inhibitor geminin act antagonistically to control the balance between proliferation and differentiation during early vertebrate eye development.

We searched for proteins that interact with Six3, a member of the Six/Sine oculis (Six/So) family, using a yeast two-hybrid screen with the entire medaka Six3 open reading frame as bait in a *Xenopus* library⁸. We identified geminin, a replication-initiation inhibitor, in this screen, and full-length medaka geminin complementary DNA (cDNA) was isolated. The predicted protein of approximately 26.5 kDa contains a highly conserved (89% similar to *Xenopus*, 87% similar to human) putative coiled-coil domain at the carboxyl terminus that also harbours the domain responsible for DNA-replication inhibition³. The overall similarity of geminin ranges from 35% (human) to 43% (*Xenopus*) at the protein level.

To assess the relevance of the interaction between geminin and Six3 we investigated the expression of their genes at the developmental stages controlled by Six3 (ref. 1). At early neurula stages (stage 16), Six3 is expressed in the anterior presumptive neuroectoderm⁹. At this time, *Gem* expression overlaps with Six3 in this region. In addition, Six3 is found in a broad domain corresponding to the future posterior neural plate (Fig. 1a, c). At the four-somite stage (stage 20), *Gem* is co-expressed with Six3 in the optic vesicles, the prosencephalon and in the overlying head ectoderm. Additionally, *Gem* is expressed in the prospective optic tectum and the midbrain–hindbrain boundary (Fig. 1b, d). Later, at organogenesis stages (stage 23 and 24), Six3 and *Gem* expression in the neuroretina overlap in proliferating marginal cells, and weak *Gem* expression is also found in the subventricular zone of the differentiating central retina and the dorsal diencephalon (black arrowhead in Fig. 1e), as well as in the cortical layer of the optic tectum (Fig. 1e–h). At the subcellular level, both proteins co-localize to the nucleus (Supplementary Fig. 1).

In the yeast two-hybrid assay, full-length Six3 specifically interacted with full-length geminin (data not shown). Six3 did not interact with chicken Six6 or mouse Six2. To confirm this specificity, we tested the interaction of Six3 and geminin in a glutathione S-transferase (GST) pull-down assay, using *in-vitro*-translated radiolabelled proteins. Geminin–GST co-precipitates medaka, *Xenopus* and human Six3 (Fig. 2a) and the closely related *Xenopus*

Six6 (Fig. 2b) but not mouse Six2 (Fig. 2a). Similarly, medaka Six3–GST strongly interacts with medaka or *Xenopus* geminin (Fig. 2c). Thus, the interaction of geminin with Six3 is evolutionarily conserved across species boundaries, and even molecules from different species interact. Related members of the Six/So gene family do not interact with geminin, however.

The two-hybrid analysis and pull-down assays independently indicate that this interaction requires the full-length geminin and Six3 proteins. Amino- or carboxy-terminal deletions of more than 30 or 75 residues, respectively, abolish the interaction of geminin with Six3, which indicates that the entire protein is required rather than a single domain (Supplementary Fig. 2).

A Six3-binding site representing the core consensus for most homeodomain-containing transcription factors¹⁰ has been identified by *in vitro* selection using the systematic evolution of ligands by exponential enrichment (SELEX) approach¹¹. We used this Six3-binding consensus sequence to perform electrophoresis mobility shift assays (EMSA). In this assay, geminin–GST does not bind to the DNA on its own. In complex with geminin–GST, Six3 still binds to DNA, resulting in a super-shifted band (Fig. 2d). This

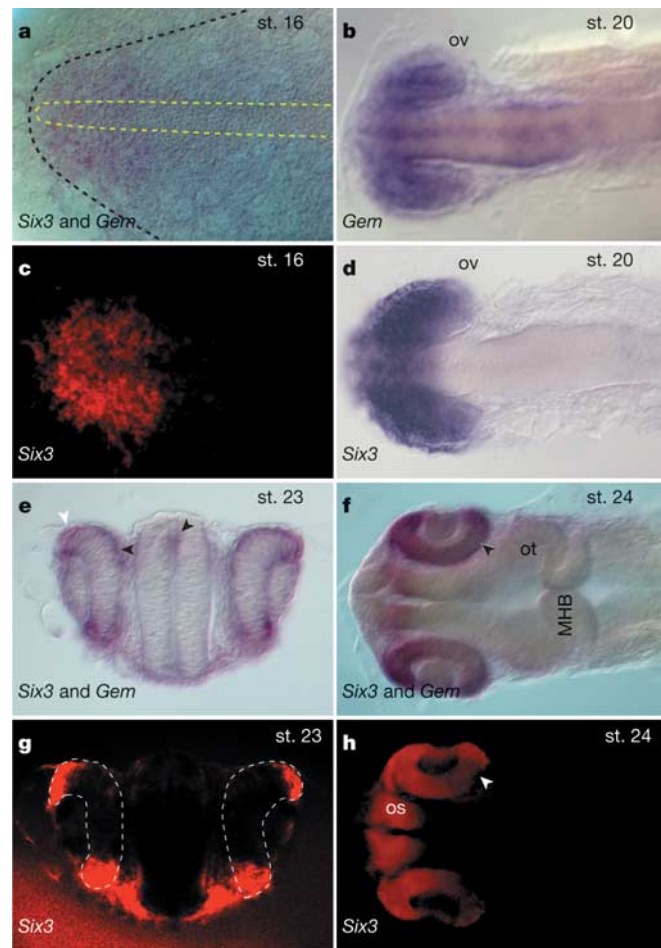


Figure 1 Spatio-temporal expression of *Gem* and *Six3*. **a–h**, *In situ* hybridization reveals that *Gem* and *Six3* are co-expressed at early stages of eye development. *Gem* (blue) and *Six3* expression patterns overlap in the anterior-most neuroectoderm at the early neurula stage (**a**, **c**), in the optic vesicles, the prosencephalon and the overlying head ectoderm at the four-somite stage (**b**, **d**), and in the neuroretina at organogenesis stages (**f**, **h**). The cross-sections show co-expression in the marginal neuroretina (white arrowhead), whereas geminin is detected in a broader domain in the central neuroretina and in the dorsal diencephalon (black arrowhead, **e**, **f**). ov, optic vesicle; os, optic stalk; MHB, midbrain–hindbrain boundary.

further corroborates the protein–protein interaction between geminin and Six3 and shows that the geminin–Six3 interaction does not affect the DNA-binding properties of Six3 in this biochemical assay.

Geminin is a negative regulator of cell-cycle progression; it inhibits the initiation of DNA replication through binding to Cdt1. This interaction prevents Cdt1 binding to the origins of replication and forming pre-replicative complexes^{4,5}. To establish whether Six3 has a direct role in cell-cycle progression, we analysed the interaction of Six3 with geminin and Cdt1. In a GST pull-down assay, Six3 (and the closely related Six6) competes with Cdt1 to bind to geminin (Fig. 2e). An increase in *Xenopus* Six3 levels specifically abolishes the interaction between *Xenopus* geminin and *Xenopus* Cdt1. This is not observed with the related *Xenopus* Six1 protein

(Fig. 2e). Conversely Cdt1 cannot displace Six3 bound to geminin, which indicates that the geminin–Six3 complex is highly stable (Supplementary Fig. 2). To discover the extent to which binding of geminin to Six3 affects its transcriptional activity, we used a cell-culture assay¹². Six3 activates reporter-gene transcription, whereas co-transfection of Six3 and *Gem* results in basal reporter-gene activity, which indicates an inhibitory effect of geminin on the transcriptional activity of Six3 (Supplementary Fig. 3). Taken together our data are consistent with Six3 and geminin having antagonistic functions in controlling cell proliferation.

To further corroborate the interaction of geminin and Six3 *in vivo*, we performed immunoprecipitation experiments in human HeLa cells. Upon co-transfection, Six3 efficiently co-immunoprecipitates geminin, and geminin efficiently co-precipitates Six3 (Fig. 2f). These findings support the co-localization results in human cells (Supplementary Fig. 1). We tested the functional significance of this interaction by overexpressing *Gem* alone or in combination with Six3 in medaka. Overexpression of *Gem* leads to severe, strictly dose-dependent phenotypes (Supplementary Table 1). At low doses, *Gem* overexpression leads to a reduction in eye size and in the most rostral forebrain (Fig. 3g, h). Intermediate doses cause severe cyclopia and loss of anterior forebrain (Fig. 3i, j), whereas at high doses the entire forebrain is lost (Fig. 3k, l), a phenotype that closely resembles that of the loss of Six3 (ref. 2). We addressed whether this is due to reduced proliferation, premature differentiation or enhanced apoptosis. Intermediately affected embryos at the late neurula stage exhibited severely reduced numbers of proliferating cells (compare Fig. 3c with d), and neuronal marker expression¹³ is induced prematurely at early neurula stages (Fig. 3e, f).

In the absence of Six3, cells that normally express Six3 failed to proliferate and instead became apoptotic². Overexpression of *Gem* leads to the same cellular consequences—increased cell death in the anterior neural plate within the Six3 expression domain (Fig. 4a, b)—in line with the proposed antagonistic interaction of Six3 and geminin, which maintains the equilibrium between proliferation and differentiation. *Gem* overexpression also resembles loss of Six3 at the molecular level. In *Gem*-injected embryos, a dose-dependent reduction of the expression of the retina-specific homeobox gene *Rx2* was found, which is consistent with the microphthalmia observed later (Supplementary Fig. 4). At higher doses, *Rx2* expression domains were fused anteriorly, which indicates the loss of proximal eye structure, resulting in cyclopia (Supplementary Fig. 4). The analysis of molecular markers specific for proximal forebrain structures¹⁴ confirmed the higher sensitivity of this region to *Gem* overexpression (Supplementary Fig. 4). Taken together, *Gem* gain-of-function phenocopies the loss of Six3 at the molecular level, which indicates an inhibitory effect of geminin on Six3 activity.

Accordingly, *Gem* gain-of-function phenotypes are efficiently rescued by Six3 in a dose-dependent manner (Supplementary Table 1). This is best explained by the two proteins directly inhibiting each other *in vivo*. To establish the direction of this inhibition (does Six3 inhibit geminin or does geminin inhibit Six3?), we used a variant of Six3 that does not bind to DNA. To generate this variant we replaced Trp 216 with Arg (Six3(W216R)). This amino-acid substitution in the DNA-recognition helix of the homeodomain abolishes DNA binding of other homeodomain-containing transcription factors such as Pit-1 or Six5 (refs 15, 16). The mutated form of Six3 does not bind to DNA (Fig. 2d). Nevertheless Six3(W216R) still interacts with geminin in the GST pull-down assay (Fig. 2a) and can efficiently compete with Cdt1 (Fig. 2e). Co-injection of Six3(W216R) efficiently rescues *Gem* overexpression phenotypes (Supplementary Table 1), which indicates that DNA binding of Six3 is dispensable for this activity. This strongly suggests that Six3 antagonizes geminin activity by direct interaction and sequestration.

Overexpression of Six3 and of the closely related Six6 induces an

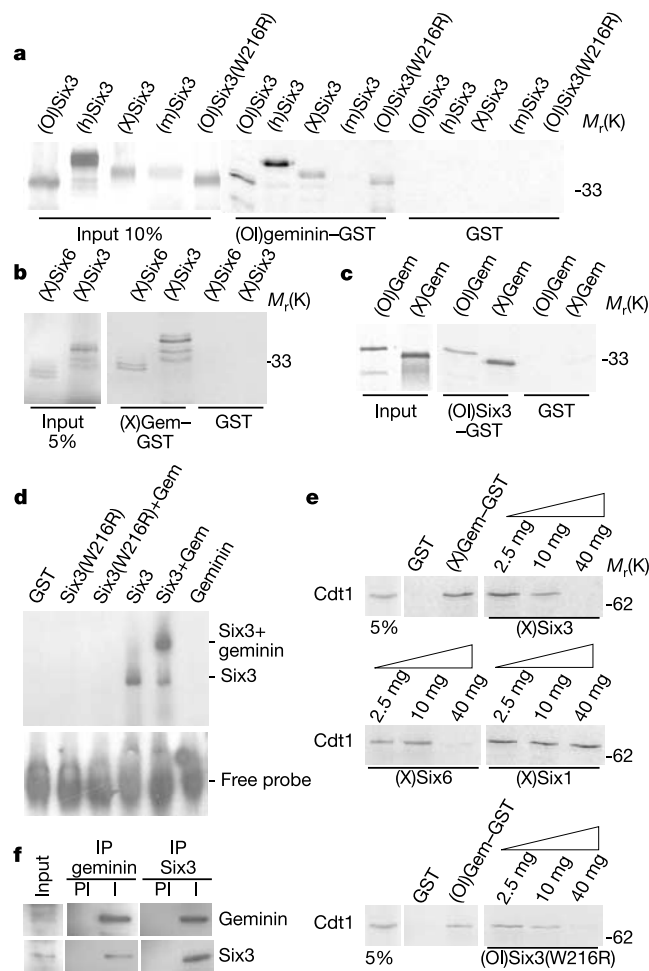


Figure 2 Biochemical analysis of the geminin–Six3 interaction. **a**, GST pull-down assay confirms the interaction of medaka geminin ((O)geminin–GST) with Six3 homologues from medaka, human and *Xenopus* ((O)Six3, (h)Six3, (X)Six3) and with the point mutant *O. latipes* Six3(W216R), but not with mouse Six2 ((m)Six2). **b**, *Xenopus* geminin–GST fusion protein interacts with *Xenopus* Six3 or the closely related *Xenopus* Six6. **c**, Medaka Six3–GST fusion protein ((O)Six3–GST) interacts with medaka and *Xenopus* geminin homologues ((O)Gem, (X)Gem). **d**, EMSA showed an upper band formed by the complex of geminin and Six3 with the DNA probe and a lower band containing Six3 bound to DNA. In this assay, geminin or Six3(W216R) alone or together with geminin did not bind to the probe. **e**, *Xenopus* Six3, Six6 and *O. latipes* Six3(W216R), but not *Xenopus* Six1, compete with Cdt1 to bind to geminin–GST fusion protein ((X)Gem–GST) in a GST pull-down assay. **f**, Immunoprecipitation of transfected HeLa cell lysate with antibodies to geminin or Six3 followed by immunoblotting with antibodies to Six3 or geminin demonstrates their interaction *in vivo*.

enlargement of the retina in a variety of vertebrate species and increases cell proliferation^{1,17}. If geminin and Six3 act antagonistically the expected consequence of loss of *Gem* function is an increase in cell proliferation. We inactivated endogenous *Gem* by morpholino-mediated gene knockdown¹⁸. Control experiments show that a morpholino oligonucleotide complementary to the *Gem* 5' untranslated region (UTR) (MoGem) specifically and stably inhibits geminin translation in a concentration-dependent manner (Supplementary Fig. 5). In the embryo, inactivation of *Gem* results in enlarged optic vesicles at early somitogenesis stages highlighted by an expansion of the *Rx2* expression domain (Fig. 4c, d). Co-staining of *En2* shows that other regions of the central nervous system, such as the midbrain–hindbrain boundary, are not affected (Fig. 4c, d).

Optic-vesicle enlargement was due to increased cell proliferation. We determined the number of mitotically active cells in the retinal primordia of MoGem-injected embryos using an antibody directed against phospho-histone H3 (H3P). To detect mitotic cells in the developing eyes, we co-stained the injected embryos with the eye-field marker *Rx3*. *Gem* inactivation increases the number of mitotic cells within the *Rx3* expression domain by 20% at the four-somite stage. In Fig. 4e, f, representative optical sections of these embryos are shown. The number of apoptotic cells (assayed by TdT-mediated dUTP nick end labelling, TUNEL) in the absence of *Gem* activity was slightly increased compared with the control in all embryos examined (16 embryos; data not shown). This excludes an anti-apoptotic mechanism as the cause of the observed optic vesicle enlargement in the absence of *Gem* function. By contrast, it indicates that the increase in mitotically active cells caused by *Gem* inactivation may be underestimated. Thus, in agreement with their proposed antagonistic activity, the loss of *Gem* function, like the gain of *Six3* function, leads to an increase in mitotically

active cells within the developing optic vesicle, and consequently to an increase in the size of the vesicles. It does not affect other structures of the embryo.

All experimental evidence so far argues for a direct, antagonistic interaction of Six3 and geminin to control cell proliferation and differentiation. To address whether they act in the same molecular pathway, we co-injected MoGem and Six3 messenger RNA at concentrations at which neither of them alone has an effect on eye development. This resulted in a marked enlargement of the optic vesicles as revealed by morphology and *Rx2* expression (24 of 53 embryos examined, Fig. 4g, h and Supplementary Fig. 6). These data consolidate the model and demonstrate that geminin and Six3 act as antagonists in the same molecular pathway, controlling the balance of cell proliferation and differentiation during early vertebrate eye development. Interestingly, this interaction is conserved among the vertebrate Six3 and geminin proteins and seems not to be limited to vertebrates. In *Drosophila*, the effect of *Gem* overexpression cannot be simply explained by an S-phase inhibitory effect¹⁹. *Optix*, the *Drosophila* orthologue of Six3 (ref. 20), which is common with its vertebrate counterpart can induce ectopic eye formation²¹, is a candidate for this interaction. The sequestration of geminin may represent a novel and widely used mechanism for a transcription factor to interfere with cell-cycle progression.

In an accompanying manuscript, Kessel and colleagues report that geminin interacts with the Hox proteins²². The complex they form does not bind to DNA, and our data indicate that the affinity of the geminin–Six3 complex for a Six3-binding consensus sequence (see above) is unaffected. Although the entire Six3 protein (including the DNA-binding Six domain) is required for the interaction with geminin, the HOX homedomain is sufficient to bind to geminin. This indicates that geminin uses a slightly different mechanism to regulate the activity of these transcription factors.

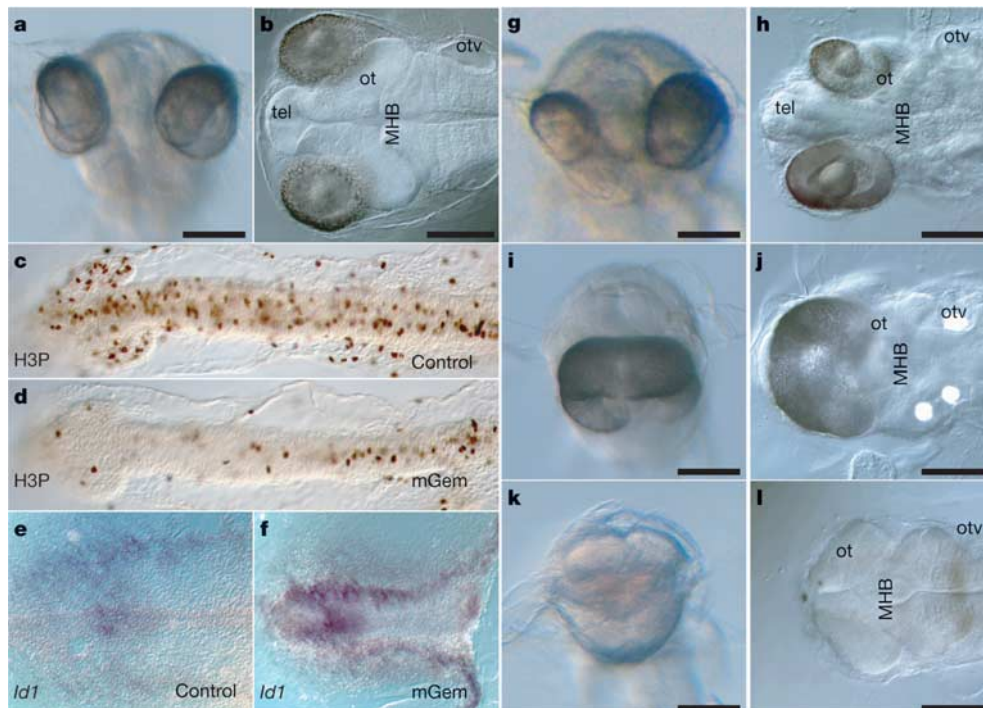


Figure 3 *Gem* overexpression results in specific eye and brain defects. Frontal (a, g, i, k) and dorsal (b, h, j, l) view of medaka embryos at 3 days of development after injection of GFP mRNA (a, b) or *Gem* mRNA (g–l). Overexpression of *Gem* leads to a reduction in the size of the eyes and the most rostral part of the forebrain at low doses (g, h). i, j, Intermediate doses cause severe cyclopia and loss of the anterior forebrain structures,

whereas at the highest doses *Gem* injection results in the complete loss of the forebrain (k, l). c, d, Geminin overexpression decreases cell proliferation, as visualized by anti-phospho-histone H3 staining. Neural marker *Id1* is prematurely induced in *Gem*-injected embryos (f) compared with control embryos (e) at late gastrula stage (stage 16). Scale bars, 100 μ m. tel, telencephalon; otv, otic vesicle; ot, optic tectum.

In addition to its role as cell-cycle inhibitor, *Gem* induces premature neuronal differentiation²³ (Fig. 3e, f), combining two distinct but related functions for nervous-system development¹⁹ in one molecule. Our results show that *Six3* and *Gem* act antagonistically in controlling proliferation and differentiation. High levels of *Six3* at the early stages of eye development inhibit *Gem* activity, which allows proliferation and prevents premature neuronal differentiation. Later, *Gem* is maintained in the differentiating central neural retina, where it does not overlap with *Six3*, to promote cell-cycle exit, a necessary prerequisite for the initiation of neuronal differentiation. Geminin and *Six3* act as a pair of antagonists: *Six3* influences the activity of geminin, and geminin influences the activity of *Six3*. The future identification and analysis of more *Six3* target sites will be a prerequisite to analyse this in a wider context. Taken together, *Six3* and *Gem* represent a clear link between cell proliferation and cell differentiation, regulating the delicate

balance that leads to the precise control of organ size during eye development. □

Methods

Medaka stocks

The Cab-strain of wild-type *O. latipes* from a closed stock at EMBL-Heidelberg was kept as described previously²⁴. Embryos were staged according to ref. 25.

Isolation of geminin cDNA

A fragment of about 100 base pairs encoding medaka *Gem* was amplified from a stage 18 cDNA library using degenerate PCR primers (up 5'-TAYTGGGAARGARGTIGCIGA; low 5'-TYNGCCATRTATGTGIACRTG). The PCR conditions used were five cycles at 94 °C for 1 min, 45 °C for 2 min and 72 °C for 4 min, followed by 30 cycles with annealing at 50 °C. The resulting fragment was cloned into the TOPO TA vector (Invitrogen), and the sequence was used to design specific PCR primers to isolate the full-length cDNA clone from the same stage-18 library using standard molecular biology techniques. The sequence has been submitted to the EMBL database (accession number AJ608707).

Whole-mount *in situ* hybridization and TUNEL assay

Whole-mount *in situ* hybridization was performed using digoxigenin-labelled RNA riboprobes as described previously⁹. Apoptotic cells were visualized as described previously¹⁴.

Cell culture immunostaining and immunoprecipitation

Human HeLa cells were grown in DMEM with 10% fetal bovine serum. Cells were transfected using FuGENE 6 (Roche) following the manufacturer's instructions. For immunofluorescence, cells were fixed in 3% paraformaldehyde for 20 min, washed twice in PBS and permeabilized with 0.1% Triton X-100 in PBS. Primary antibodies anti-human *Six3* (mouse, a gift from P. Bovolenta) and anti-human geminin (rabbit, Santa Cruz) were used at 1:1,000 and 1:200 dilution, respectively. Alexa-488-conjugated anti-mouse and biotinylated anti-rabbit, followed by incubation with Alexa-568-conjugated streptavidin, were used as secondary antibodies. For immunoprecipitation, transfected cells were lysed in a solution containing 150 mM NaCl, 1% Nonidet P-40, 50 mM Tris-HCl pH 8 and protease inhibitors. One milligram of total proteins was used per immunoprecipitation with 5 µl of anti-*Six3* antibody or 10 µg of anti-geminin antibody. After five washes in PBS 0.1% Nonidet P-40, samples were subjected to SDS-polyacrylamide gel electrophoresis (PAGE) and blotted onto PVDF membrane for western-blot analysis.

GST pull-down assay and EMSA

The GST pull-down assay was performed as described previously⁸, with minor modifications. Briefly, after purification with glutathione sepharose beads, GST and GST fusion protein were loaded on a SDS-polyacrylamide gel for quantification. Two micrograms of protein coupled with the matrix were used in each sample. *Xenopus* *Six3* and *Six1* proteins were expressed and purified as GST fusion proteins in the pGEX-KG bacteria vector. They were eluted from the matrix by enzymatic cleavage adding 1% (w/w) thrombin (human, Boehringer Mannheim) to the estimated amount of fusion protein bound to the beads in thrombin cleavage buffer (50 mM Tris-Cl, pH 7.5; 150 mM NaCl; 2.5 mM CaCl₂). The reaction was allowed to proceed for 2 h at room temperature. Cleaved purified protein was recovered in the supernatant and quantified by SDS-PAGE. *Xenopus* *Six3*, *Six6*, *O. latipes* *Six3* (W216R) or *Xenopus* *Six1* (2.5 µg, 10 µg or 40 µg) were added to the standard GST pull-down assay with *Xenopus* geminin-GST and *in-vitro*-translated *Xenopus* Cdt1 (³⁵S-labelled), in the immunoprecipitation buffer⁸ for 2 h at 4 °C.

EMSA was performed as described previously¹¹ with the following modifications. The DNA double-stranded probe (GGAGTCCGTGGGGGATGTGAGATGGATTAAATAGCTGTAGCGTTATTGG) was end-labelled using T4 polynucleotide kinase with [γ-³²P]dATP. One hundred nanograms of total bacterial lysate of *Escherichia coli* strain BL21 (DE3) expressing GST or GST-fusion protein was added to the probe. Electrophoresis was performed at 200 V at 4 °C for 4–6 h.

RNA injections

RNA injections were performed as described previously¹. We tested the effects of *Gem* overexpression at different concentrations. Twenty-five nanograms per microlitre of *Gem* mRNA injection resulted in 11% of embryos mildly affected (Fig. 3g, h), 1% cyclopic (Fig. 3i, j) and 1% severely affected (Fig. 3k, l) (*n* = 174). Fifty nanograms per microlitre resulted in 26%, 9% and 4%, respectively (*n* = 218). For rescue experiments, 50 and 100 ng µl⁻¹ of *Six3* were co-injected with geminin mRNA. Control experiments were performed by injecting an equal amount of GFP mRNA, which had no effect on embryological development.

Morpholino injections and control experiments

Morpholino (Gene Tools) targeted against *Gem* (MoGem, TTTTCCTTTACAATGTGTAGCCCGC) was injected, and control experiments were performed as described previously². Control injections were performed using GFP morpholino (MoGFP) or a geminin morpholino containing a five-base mismatch (MissMo, TTTaCCTTaACAATcTGTAtCCtGC). MoGem was used at a final concentration of 0.6 mM, which resulted in 47% of embryos being affected (46 of 98). For co-injection experiments, 0.06 mM MoGem was injected with 10 ng µl⁻¹ of *Six3* or GFP mRNA resulting in 24 of 53 and 1 of 38 embryos, respectively, with expanded *Rx2* expression.

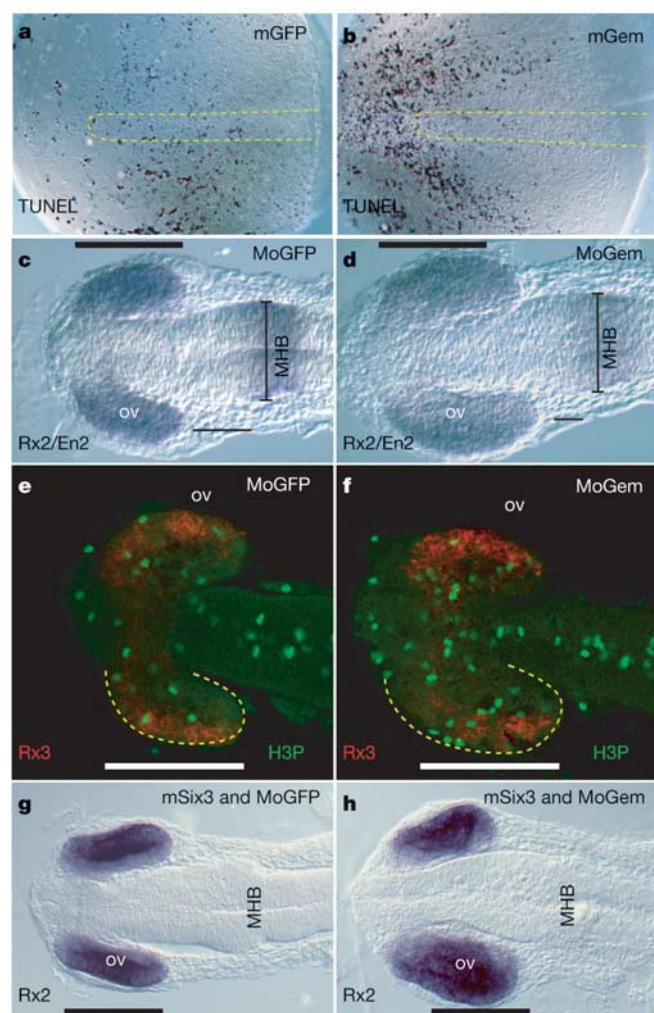


Figure 4 *Gem* overexpression results in ectopic cell death, whereas geminin inactivation increases cell proliferation and induces retina enlargement. **a, b**, Ectopic cell death (detected by TUNEL assay) in response to *Gem* overexpression (**a, b**) overlaps with the *Six3* expression in the anterior neuroectoderm. **c, d**, Geminin morpholino injection (0.6 mM) results in enlarged optic vesicles revealed by expression of *Rx2*. Scale bars indicate that the midbrain-hindbrain boundary expressing *En2* is unaffected. **e, f**, There are an increased number of H3P-positive cells (in green, **f**) in the eye field (*Rx3* expression, in red) compared with control embryos injected with MoGFP (0.6 mM) (**e**). **g, h**, *Six3* mRNA (10 ng µl⁻¹) and MoGem (0.06 mM) injection enlarges *Rx2* expression domain (**h**) compared with controls injected with *Six3* mRNA (10 ng µl⁻¹) and MoGFP (0.06 mM) (**g**). Scale bars, 100 µm.

Whole-mount antibody staining

After *in situ* detection of Rx3 expression using Fast Red (Boehringer Mannheim) as the fluorescent substrate, phosphorylation of histone H3 at Ser 10 was revealed using a polyclonal antibody (Upstate Biotechnology, 1:1,000 dilution). Secondary anti-rabbit antibody fluorescein-conjugate was used, and embryos were analysed using a confocal microscope (Leica TCS-SP). Optical sections of 4 μ m were recorded, and positive cells were counted. Alternatively anti-phospho-histone H3 antibody was detected with a secondary peroxidase-conjugated anti-rabbit antibody, followed by DAB staining.

Received 1 October; accepted 11 December 2003; doi:10.1038/nature02292.

- Loosli, F., Winkler, S. & Wittbrodt, J. Six3 overexpression initiates the formation of ectopic retina. *Genes Dev.* **13**, 649–654 (1999).
- Carl, M., Loosli, F. & Wittbrodt, J. Six3 inactivation reveals its essential role for the formation and patterning of the vertebrate eye. *Development* **129**, 4057–4063 (2002).
- McGarry, T. J. & Kirschner, M. W. Geminin, an inhibitor of DNA replication, is degraded during mitosis. *Cell* **93**, 1043–1053 (1998).
- Wohlschlegel, J. A. *et al.* Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. *Science* **290**, 2309–2312 (2000).
- Tada, S., Li, A., Maiorano, D., Mechali, M. & Blow, J. J. Repression of origin assembly in metaphase depends on inhibition of RLF-B/Cdt1 by geminin. *Nature Cell Biol.* **3**, 107–113 (2001).
- Maiorano, D., Moreau, J. & Mechali, M. XCDT1 is required for the assembly of pre-replicative complexes in *Xenopus laevis*. *Nature* **404**, 622–625 (2000).
- Lagutin, O. V. *et al.* Six3 repression of Wnt signaling in the anterior neuroectoderm is essential for vertebrate forebrain development. *Genes Dev.* **17**, 368–379 (2003).
- Tessmar, K., Loosli, F. & Wittbrodt, J. A screen for co-factors of Six3. *Mech. Dev.* **117**, 103–113 (2002).
- Loosli, F., Köster, R. W., Carl, M., Krone, A. & Wittbrodt, J. Six3, a medaka homologue of the *Drosophila* homeobox gene sine oculis is expressed in the anterior embryonic shield and the developing eye. *Mech. Dev.* **74**, 159–164 (1998).
- Treisman, J., Harris, E., Wilson, D. & Desplan, C. The homeodomain: a new face for the helix-turn-helix? *Bioessays* **14**, 145–150 (1992).
- Zhu, C. C. *et al.* Six3-mediated auto repression and eye development requires its interaction with members of the Groucho-related family of co-repressors. *Development* **129**, 2835–2849 (2002).
- Goudreau, G. *et al.* Mutually regulated expression of Pax6 and Six3 and its implications for the Pax6 haploinsufficient lens phenotype. *Proc. Natl Acad. Sci. USA* **99**, 8719–8724 (2002).
- Sawai, S. & Campos-Ortega, J. A. A zebrafish Id homologue and its pattern of expression during embryogenesis. *Mech. Dev.* **65**, 175–185 (1997).
- Winkler, S., Loosli, F., Henrich, T., Wakamatsu, Y. & Wittbrodt, J. The conditional medaka mutation eyeless uncouples patterning and morphogenesis of the eye. *Development* **127**, 1911–1919 (2000).
- Liang, J., Moye-Rowley, S. & Maurer, R. A. *In vivo* mutational analysis of the DNA binding domain of the tissue-specific transcription factor, Pit-1. *J. Biol. Chem.* **270**, 25520–24425 (1995).
- Sato, S. *et al.* Identification of transcriptional targets for Six5: implication for the pathogenesis of myotonic dystrophy type 1. *Hum. Mol. Genet.* **11**, 1045–1058 (2002).
- Zuber, M. E., Perron, M., Philpott, A., Bang, A. & Harris, W. A. Giant eyes in *Xenopus laevis* by overexpression of XOptx2. *Cell* **98**, 341–352 (1999).
- Nasevicius, A. & Ekker, S. C. Effective targeted gene 'knockdown' in zebrafish. *Nature Genet.* **26**, 216–220 (2000).
- Quinn, L. M., Herr, A., McGarry, T. J. & Richardson, H. The *Drosophila* Geminin homolog: roles for Geminin in limiting DNA replication, in anaphase and in neurogenesis. *Genes Dev.* **15**, 2741–2754 (2001).
- Toy, J., Yang, J.-M., Leppert, G. & Sundin, O. H. The *Optx2* homeobox gene is expressed in early precursors of the eye and activates retina-specific genes. *Proc. Natl Acad. Sci. USA* **95**, 10643–10648 (1998).
- Seimiya, M. & Gehring, W. J. The *Drosophila* homeobox gene optix is capable of inducing ectopic eyes by an eyeless-independent mechanism. *Development* **127**, 1879–1886 (2000).
- Luo, L., Yang, X., Takiyama, Y., Knoetgen, H. & Kessel, M. The cell-cycle regulator geminin inhibits Hox function through direct and polycomb-mediated interactions. *Nature* **427**, 749–753 (2004).
- Kroll, K. L., Salic, A. N., Evans, L. M. & Kirschner, M. W. Geminin, a neuralizing molecule that demarcates the future neural plate at the onset of gastrulation. *Development* **125**, 3247–3258 (1998).
- Köster, R., Stick, R., Loosli, F. & Wittbrodt, J. Medaka spalt acts as a target gene of hedgehog signaling. *Development* **124**, 3147–3156 (1997).
- Iwamatsu, T. Stages of normal development in the medaka *Oryzias latipes*. *Zoo. Sci.* **11**, 825–839 (1994).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We acknowledge F. Loosli and J. Martinez for their critical input and discussion; P. Bovolenta and J. Lopez-Rios for the human Six3 construct and antibody; Z. Lygerou, T. Nishimoto, A. Mansouri, A. Nebreda and M. Mechali for providing plasmids and cells; and A. Akhtar and V. Neubrand for methodological advice. We thank C. Grabher, M. Rembold and W. Norton for critically reading the manuscript, E. Grzebisz for fish husbandry and A. Krone for technical assistance. We would like to thank M. Kessel for communicating unpublished results and for critical input. This work was supported by grants from the Deutsche Forschungsgemeinschaft, Collaborative Research Centre 488, the EU and HFSP to J.W.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.W. (Jochen.Wittbrodt@EMBL.de). The geminin sequence has been submitted to the EMBL database (accession number AJ608707); whole-mount expression data have been submitted to MEPD (OL_emx1, OL_geminin, OL_rx2, OL_six3, OL_vax1; <http://www.embl-heidelberg.de/mepd/>).

The cell-cycle regulator geminin inhibits Hox function through direct and polycomb-mediated interactions

Lingfei Luo¹, Xiaoping Yang², Yoshihiro Takiyama³, Hendrik Knoetgen¹ & Michael Kessel¹

¹Research Group Developmental Biology, Department of Molecular Cell Biology, and ²Department of Molecular Genetics, Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany

³Department of Stem Cell Biology, Research Institute for Radiation Biology and Medicine, Hiroshima University, Hiroshima 734-8553, Japan

Embryonic development is tightly controlled. The clustered genes of the *Hox* family of homeobox proteins play an important part in regulating this development and also proliferation. They specify embryonic structures along the body axis, and are associated with normal and malignant cell growth^{1–4}. The cell-cycle regulator geminin controls replication by binding to the licensing factor Cdt1, and is involved in neural differentiation^{5–7}. Here, we show that murine geminin associates transiently with members of the *Hox*-repressing polycomb complex, with the chromatin of *Hox* regulatory DNA elements and with *Hox* proteins. Gain- and loss-of-function experiments in the chick neural tube demonstrate that geminin modulates the anterior boundary of *Hoxb9* transcription, which suggests a polycomb-like activity for geminin. The interaction between geminin and *Hox* proteins prevents *Hox* proteins from binding to DNA, inhibits *Hox*-dependent transcriptional activation of reporter and endogenous downstream target genes, and displaces Cdt1 from its complex with geminin. By establishing competitive regulation, geminin functions as a coordinator of developmental and proliferative control.

To identify proteins that interact with geminin during embryogenesis, we performed a two-hybrid screen using a complementary DNA library prepared from 8.5 days post coitum (d.p.c.) mouse embryos. Eight positive yeast clones were shown to synthesize geminin-binding proteins (Fig. 1a). Three independent cDNAs each encoded parts of the homeodomain proteins Hoxd10 and Hoxa11. One clone encoded the sex comb on midleg homologue 1

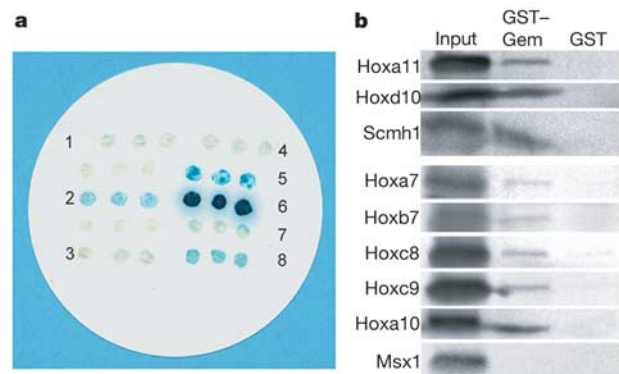


Figure 1 Isolation of Hox and Scmh1 as geminin-binding proteins. **a**, Yeast clones selected by two-hybrid analysis. Eight positive clones were shown to encode geminin-binding proteins, as visualized by β -galactosidase activity. Clones 1, 2, 8 encode Hoxa11, and clones 3, 4, 7 encode Hoxd10. Clone 6 encodes Scmh1. **b**, Pull-down assays. All the full-length *in-vitro*-transcription/translation products of the indicated genes except Msx1 were specifically bound by a GST-geminin fusion protein but not by GST alone.

(Scmh1) protein, which is the mouse homologue of *Drosophila* Scm, a member of the polycomb multiprotein complex^{8,9}. Full-length Hox, mesenchymal transcription factor Msx1 and Scmh1 proteins were synthesized *in vitro* in the presence of [³⁵S]-methionine. These radiolabelled proteins were tested for their ability to bind to a recombinant glutathione S-transferase (GST)–geminin fusion protein and to GST alone, which acted as the control (Fig. 1b). In such pull-down assays, all the Hox proteins (Hoxa11, Hoxd10, Hoxa7, Hoxb7, Hoxc8, Hoxc9 and Hoxa10) and Scmh1, but not full-length Msx1, tested bound directly to GST–geminin. They exhibited no appreciable binding to GST alone. Together, our results identify homeodomain proteins of the Hox family, as well as the polycomb group member Scmh1, as binding partners of the cell-cycle regulator geminin.

To test whether geminin associates with other polycomb complex members, we analysed its *in vivo* co-localizations with Rae28 (also known as Mph1)¹⁰ and Mel18 (ref. 11) proteins. Double immunofluorescence staining of U2-OS cells was performed with anti-geminin antibodies and antibodies against Rae28 or Mel18 (Supplementary Fig. 1). Endogenous geminin co-localized extensively with Rae28 and Mel18 in the nuclei of the cultured cells. The two polycomb proteins were detected in the nuclei of all cells; however, several nuclei did not contain geminin. A close inspection of these cells revealed that the presence of geminin depended on the phase of the cell cycle (Fig. 2a). Geminin was not detectable during interphase; it accumulated in the nucleus and persisted throughout mitosis until the anaphase–telophase transition. By contrast, Rae28 and Mel18 were continuously expressed throughout the cell cycle. Direct evidence for an *in vivo* association between geminin and a polycomb protein was obtained by immunoprecipitation from 11.5-d.p.c. mouse embryonic extracts using anti-geminin antibodies. A pre-immune rabbit serum was used as a negative control. An immunoprecipitated band with a mobility of 120 kDa was recognized by anti-Rae28 antibodies⁹ (Fig. 2b), which indicates

that the protein complex isolated from mouse embryos containing geminin also contained Rae28. These data demonstrate a cell-cycle-dependent association of geminin and the polycomb complex protein Rae28.

A fraction of geminin is chromatin associated¹². To investigate whether geminin associates with *Hox* regulatory elements *in vivo*, we performed chromatin immunoprecipitation (ChIP) assays to isolate DNA fragments that were bound to geminin-including protein complexes in primary cultured mouse embryonic fibroblasts (MEFs). Four *Hox* regulatory elements within the *Hoxd11* gene bind to promyelocytic leukaemia zinc finger (Plzf), a protein that associates with polycomb complex members and mediates transcriptional repression of *Hox* genes¹³. Three of these Plzf-binding sites, located within the *Hoxd11* intron or 3' untranslated region (UTR), were co-precipitated by geminin antibodies but not by pre-immune serum (Fig. 2c). These results demonstrate that geminin associates *in vivo* with the *Hox* regulatory DNA elements that anchor Plzf.

To analyse the effect of ectopically expressed geminin on *Hox* gene transcription, we overexpressed geminin in the neural tube of chick embryos using *in ovo* electroporation. *CMV-EGFP-Gem* plus *CMV-Gem* or the control vector *CMV-EGFP* were injected into the neural tube of HH9–11-stage chick embryos, that is, at an embryonic time when *Abd-B*-related *Hox* genes such as *Hoxb9* become activated in the posterior body region. The plasmids were electroporated into the right side of the neural tube, where green fluorescent protein (GFP) expression was confirmed after 24 h of incubation and recorded (Fig. 3a, b). The embryos were then incubated for another 24 h, fixed at stage HH18–19, and examined using whole-mount *in situ* hybridization. The *Hoxb9* anterior transcription boundary was posteriorly shifted by the length of one to two somites on the electroporated, right side (Fig. 3c). By contrast, there was no posterior shift of the endogenous *Hoxb9* transcription domain in control embryos with only *CMV-EGFP*

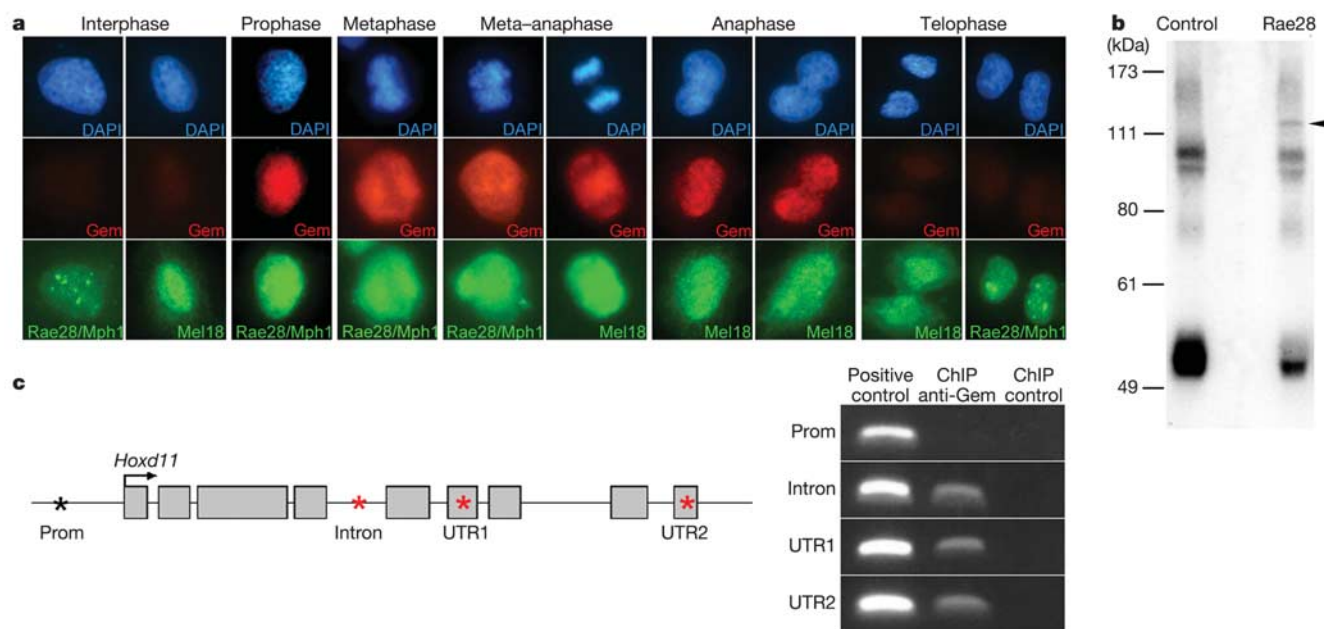


Figure 2 Geminin associates with the polycomb complex and *Hox* regulatory DNA elements *in vivo*. **a**, Subcellular co-localization of geminin and Rae28 or Mel18 during different phases of mitosis. Note the absence of geminin during interphase and telophase. DAPI, 4,6-diamidino-2-phenylindole. **b**, Co-immunoprecipitation of geminin and Rae28 from 11.5-d.p.c. mouse embryonic extracts. The arrowhead indicates that Rae28 was

specifically co-precipitated by anti-geminin antibodies. **c**, Three Plzf-binding *Hox* regulatory elements within the *Hoxd11* intron and 3' UTR (red asterisks) were identified by ChIP analysis. The *Hox* regulatory element in the promoter region (black asterisk) was not detected.

electroporated (Fig. 3d). This result shows that overexpressed geminin inhibits *Hox* gene activation.

To better characterize the role of geminin in the polycomb complex, we determined the geminin-binding domains of Scmh1. Binding of recombinant, histidine-tagged geminin (His-geminin) to an array of Scmh1 peptides revealed the domain of Scmh1 that was rich in basic amino acids to be the geminin-binding region, which lies outside the SPM domain (a conserved domain shared by Scm and polyhomeotic proteins), through which Scmh1 associates with other polycomb members (Fig. 3e)⁹. Using this information, a dominant-negative form of Scmh1 (amino acids 508–585, dnScmh1), which included the geminin-binding domain but not the SPM domain, was designed. The binding of dnScmh1 to geminin was confirmed using a pull-down assay (Fig. 3f). *In ovo* co-electroporations of a *dnScmh1* expression vector and *CMV*–

EGFP to the right side of the neural tube and whole-mount *in situ* hybridization to *Hoxb9* (Fig. 3g–j) were performed as described above. We observed a derepression of *Hoxb9* transcription, one somite length anterior of the normal expression boundary (Fig. 3i). Direct elimination of geminin was carried out by the co-electroporation of small interfering RNA (siRNA) against the chick geminin gene (siGem) and *CMV*–*EGFP*. siRNA against luciferase was used as the control (siLuc; Fig. 3k–n)¹⁴. A pronounced derepression of *Hoxb9* by the length of one-and-a-half to two somites was observed (Fig. 3m). These findings show that the inhibition of *Hox* gene transcription by geminin is due to a geminin–polycomb interaction, that is, that geminin behaves like a polycomb protein *in vivo*.

We identified the homeodomains of both Hoxa11 (Fig. 4a, black frame) and Hoxa7 (data not shown) as geminin-binding regions, with two clusters of basic amino acids as the core binding sequences, through the application of peptide arrays. The Hoxa11 interaction domain of geminin overlaps partly with the coiled-coil domain, which binds to Cdt1 (data not shown)^{5,15}. These findings suggest that geminin is a specific antagonist of DNA binding by the Hox homeodomains. We performed electrophoretic mobility shift assays (EMSA), applying *in-vitro*-transcribed/translated Hoxd10, Hoxa11, Hoxb7 and Msx1 proteins, double-stranded oligonucleotides (including their respective consensus-binding sequences) and the recombinant His-geminin protein. The translated Hox proteins produced prominent shifts in the size of oligonucleotide bands during electrophoresis (Fig. 4b, lanes 2, 4, 6 and 9). Pre-incubations of geminin with Hox proteins resulted in the release of free probe, hence the significant reduction of the shifted bands (Fig. 4b, lanes 3, 5, 7 and 10). By contrast, the binding of Msx1 to its target sequence was not attenuated by geminin (Fig. 4b, lanes 11 and 12). In summary, geminin inhibits the binding of Hox proteins to their target DNA sequences as a result of interacting with, and thus blocking, their homeodomains. The amino terminus of Hoxa11 had a small positive influence on geminin binding, whereas N-terminal Msx1 sequences inhibited the interaction of geminin with the Msx1 homeodomain, as indicated by LacZ activities in two-hybrid and binding studies with truncated proteins (data not shown). From these results we conclude that the N terminus influences the binding of geminin to a homeodomain protein. In an accompanying paper, Wittbrodt and colleagues show that the entire Six3 homeoprotein is required for an interaction with geminin, and that geminin super-shifts, but not attenuates, a Six3–DNA complex¹⁶.

To examine whether geminin inhibits the transcriptional activation promoted by Hox *in vivo*, we designed a luciferase-reporter construct with a triple Hoxa11-binding sequence. Expression of luciferase was enhanced tenfold by the overexpression of Hoxa11. This increased level was reduced by 60% if, in addition to Hoxa11, geminin was overexpressed (Supplementary Fig. 2). Similarly, when a 500-base-pair fibroblast growth factor 2 (FGF2) promoter fragment including a Hoxb7-binding site¹⁷ was applied, luciferase activity was increased markedly by overexpressed Hoxb7 in HeLa cells. This increase was reduced by 40% by geminin (Supplementary Fig. 2). Together, these data show that the geminin–Hox interaction interferes with the role of Hox proteins as transcriptional activators.

FGF2 is a downstream target gene of Hoxb7 in the melanoma cell line A375 (ref. 17). We used this well defined system to study the influence of geminin on the function of Hox proteins. *CMV*–*Gem*, a siRNA targeting human geminin messenger RNA (sihGem), or siLuc were transfected into cultured A375 cells. Subsequently, geminin, FGF2 or vimentin levels were detected by western blotting (Fig. 4c). We observed a decrease in the level of FGF2 in parallel with an increase in the level of geminin, and an increase in FGF2 in parallel with the suppression of geminin caused by specific siRNA. These results suggest that geminin modulates the function of the Hoxb7 protein *in vivo*, as detected here by measuring the product of its direct target FGF2. In addition, the FGF2 promoter region could

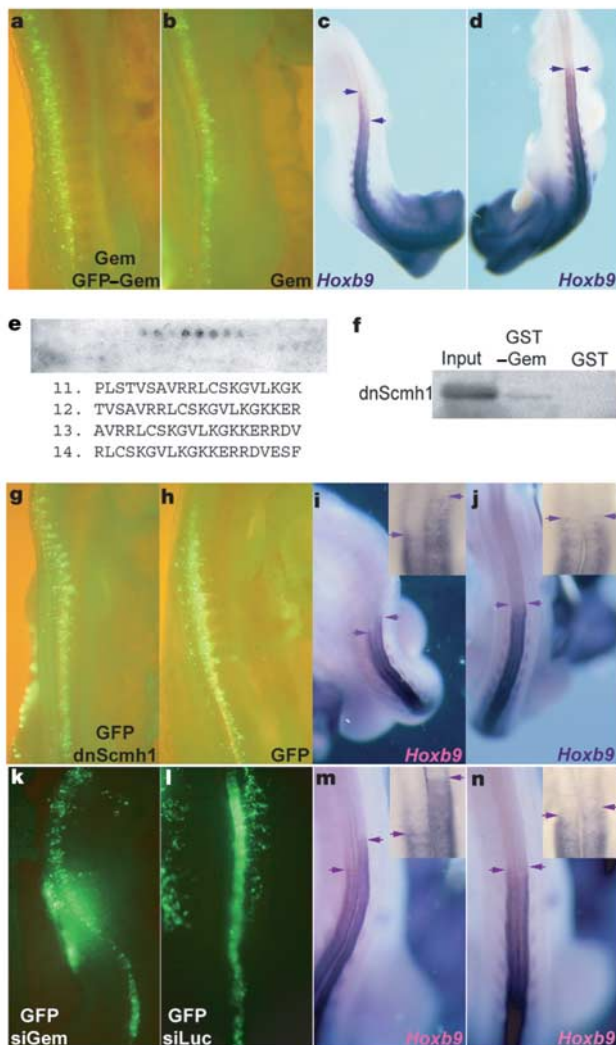


Figure 3 Geminin modulates the anterior boundary of endogenous *Hoxb9* transcription in the avian neural tube. **a–d**, The *Hoxb9* anterior transcription boundary was posteriorly shifted by ectopically expressed geminin. **e**, Binding of His-tagged geminin to a Scmh1 peptide array. Below, His-geminin-bound peptides are listed (numbered 11 to 14). Amino acids 540–568 of Scmh1 comprise the geminin-binding domain. **f**, Confirmation of dnScmh1–geminin binding through a GST–geminin pull-down assay. **g–j**, The *Hoxb9* anterior transcription boundary was anteriorly shifted by ectopically expressed dnScmh1. **k–n**, The *Hoxb9* anterior transcription boundary was anteriorly shifted by siRNA against geminin. Arrowheads indicate the anterior transcription boundary of *Hoxb9*.

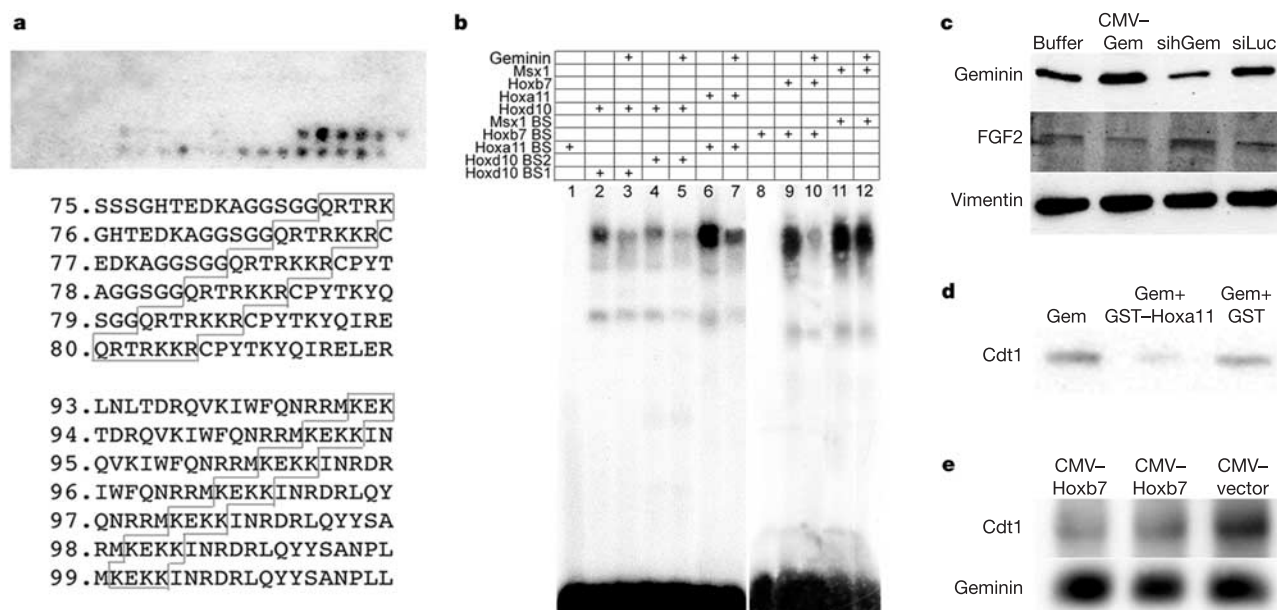


Figure 4 The geminin-Hox interaction inhibits DNA binding and competes with Cdt1-geminin complex formation. **a**, Binding of His-geminin to a Hoxa11 peptide array. Bound peptides are listed, and the common core sequences are framed in black. **b**, Geminin interferes with the binding of Hox proteins to specific double-stranded DNAs. Five different binding sites (BS) were tested in EMSAs. **c**, Geminin inhibits expression of FGF2, the downstream target of Hoxb7, in A375 cells. Note that overexpression or siRNA downregulation of geminin leads to reciprocal changes in FGF2 levels. Vimentin served as

an internal control. **d**, Competition for geminin binding between Hoxa11 and Cdt1 *in vitro*. Note that the binding of endogenous Cdt1 was significantly reduced when GST-Hoxa11, but not GST alone, was pre-bound to immobilized geminin. **e**, Competition for geminin binding between Hox proteins and Cdt1 *in vivo*. Note that the co-precipitated endogenous Cdt1 was significantly reduced when Hoxb7 or Hoxa11 was overexpressed in the cells, whereas the level of geminin itself was not affected.

not be detected when a ChIP assay was performed using an A375 cell extract and geminin antibodies (data not shown), which indicates that geminin is not recruited with Hoxb7 to its DNA target.

As both Hox and Cdt1 proteins bind to the coiled-coil domain of geminin, we further investigated whether Hox proteins compete for the binding of Cdt1 to geminin. The Cdt1 protein in the 11.5-d.p.c. embryonic extracts was pulled-down using geminin-coupled beads. The pull-down of Cdt1 was significantly decreased by pre-incubation of the geminin-coupled beads with GST-Hoxa11 recombinant protein but not with GST alone (Fig. 4d). This result was supported *in vivo* by the following results from primary cultured MEF cells. When Hoxb7 or Hoxa11 was overexpressed, the amount of Cdt1 co-precipitated by geminin antibodies was significantly reduced, in contrast to the control cells transfected with empty vector. The level of geminin itself was not affected (Fig. 4e). Together, these data indicate that Hox proteins can displace geminin from the Cdt1-geminin complex.

Geminin mRNA is widely expressed during vertebrate embryogenesis, and is expressed at slightly higher levels in neural tissues, the tail and the limb bud regions (ref. 7, and data not shown). Our study suggests that geminin is involved in two processes controlled by multiprotein complexes. One is the replication initiation of DNA during the cell cycle, which is controlled by the origin of replication complex, the minichromosome maintenance complex, as well as Cdt1 (refs 5, 6). The other is the specification of cellular identity during embryogenesis, which is controlled by Hox proteins and the polycomb complex, including Scmh1 (refs 1–3). Geminin can associate with components of both multiprotein machineries, remove them from their context, and inhibit their function, for example, the licensing of replication in the case of Cdt1 or the maintenance of Hox repression by the polycomb complex. The level as well as the timing during the cell cycle of geminin compared with proteins such as Cdt1, Hox or Scmh1 would influence the avail-

ability of a component. This conclusion is in line with the evidence for multiple roles of Hox as well as polycomb proteins not only in embryogenesis but also in proliferation, cell-cycle control and transformation^{4,18–23}. By establishing competitive regulation, geminin functions as a coordinator of developmental and proliferative control. □

Methods

Yeast two-hybrid and GST pull-down assays

The yeast two-hybrid screen was performed using the ProQuest two-hybrid system (GIBCO BRL) according to the manufacturer's instructions. A total of 177 independent colonies were selected on plates lacking leucine, tryptophan and histidine, and supplemented with 50 mM 3-aminotriazole. All these clones were assayed for β -galactosidase activity, and eight positive cDNA clones were shown to encode geminin-binding proteins. The full-length cDNAs were subcloned, and GST pull-down assays were performed as described previously²⁴.

Cells, immunostaining, immunoprecipitation and ChIP assays

U2-OS cells were cultured using the protocol described by the supplier (ATCC). The cells were fixed, permeabilized, stained for immunofluorescence and analysed as described previously²⁵.

11.5-d.p.c. mouse embryonic extracts were prepared by solubilizing the embryos in lysis buffer (20 mM Tris-HCl, pH 7.5, 100 mM KCl, 2 mM EDTA, 1% Triton X-100, 1 mM β -mercaptoethanol). One milligram of anti-geminin antibodies were pre-bound and crosslinked to 40 mg protein A Sepharose CL-4B (Amersham Pharmacia). These antibody-coupled beads were incubated with 1.5 ml mouse embryonic extracts (40 mg proteins ml⁻¹) for 2 h at 4 °C, then washed three times with washing buffer (20 mM Tris-HCl, pH 7.5, 150 mM KCl, 1 mM EDTA, 1 mM β -mercaptoethanol, 0.1% Triton X-100). The precipitated materials were eluted, analysed on a 10% SDS-polyacrylamide gel, and detected with anti-Rae28 antibodies⁹ by western blotting. For the ChIP assay, 150 μ g of geminin antibodies, 2 $\times$ 10⁶ of primary cultured MEFs and four pairs of primers (Prom: 5'-CACGAGATTGCTCAGGGCTTAG-3', 5'-CAATACTCA GCCAGCGTGGAAC-3'; Intron: 5'-TTCAGAGCCTGCTTGGCATC-3', 5'-CACTCT GGCACCTAGGCTAG-3'; UTR1: 5'-CCACTACAGCCTGAGGAAGAG-3', 5'-GACAG TGACTATGCCCAAG-3'; UTR2: 5'-CATAAGATGCACAGCAGCTCATGC-3', 5'-GTGGGTCTGGATGTATGAGCCTG-3') were applied using the ChIP assay kit (Upstate), and genomic PCR and pre-immune serum were used as controls.

In ovo electroporation

In ovo plasmid and siRNA electroporations of HH9–11-stage chick embryos and the analysis of endogenous *Hoxb9* transcription by whole-mount *in situ* hybridization were carried out as described previously^{14,24,26}.

Peptide array, EMSA and cell-transfection assays

Peptide-array membranes (Jerini) were incubated overnight with 1 µg ml⁻¹ His-geminin recombinant protein in a buffer containing 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 0.05% Tween 20 at 4 °C. Then, the His-geminin binding peptides were detected with anti-His antibodies (Novagen) using the manufacturer's protocol.

Sense strands of complementary oligonucleotides containing binding sites for Hoxa11, Hoxd10, Hoxb7 or Mx1 (refs 17, 27, 28) were end-labelled, purified and annealed with their respective antisense oligonucleotides to generate double-stranded DNA probes. *In-vitro*-transcribed/translated proteins were incubated on ice with 20,000 c.p.m. DNA probes in the retardation buffer (20 mM HEPES, pH 7.6, 4% Ficoll, 5 mM MgCl₂, 40 mM NaCl, 0.1 mM EDTA, 0.5 mM DTT, 0.2 µg µl⁻¹ poly[dI-dC]) for 1 h, with or without preincubation with 5 µg His-geminin recombinant protein. The reaction mixtures were applied to a 10% polyacrylamide gel and visualized by exposure to a BioMax film (Kodak) overnight.

Transfection of A375 cells with *CMV-Gem* or siRNAs and detections with antibodies against geminin, FGF2 (Chemicon) or vimentin were performed as described previously²⁹.

Competition assays

A total of 200 µg His-geminin recombinant protein was N-terminally coupled and crosslinked to 100 µl beads using AminoLink Plus Coupling Gel (Pierce) using the manufacturer's protocol. These geminin-coupled beads were then pre-incubated with 500 µg GST-Hoxa11 or pure GST in 20 mM Tris-HCl, pH 7.5, 50 mM KCl, 1 mM EDTA, 1 mM β-mecaptoethanol, 0.1% NP-40 for 90 min at 4 °C, followed by washing twice with the same buffer. These beads, with or without pre-incubation, were sequentially incubated with 1 ml 11.5-d.p.c. mouse embryonic extracts for 2 h at 4 °C, washed twice with the same buffer, eluted and loaded on a 10% SDS-polyacrylamide gel. The materials pulled-down by the geminin-coupled beads were detected with anti-Cdt1 antibodies⁵ by western blotting. Transfections of *CMV-Hoxb7* or *CMV-Hoxa11* into MEFs and immunoprecipitation were performed as described above.

Received 9 October; accepted 15 December 2003; doi:10.1038/nature02305.

1. Krumlauf, R. *Hox* genes in vertebrate development. *Cell* **78**, 191–200 (1994).
2. Kessel, M. & Gruss, P. Homeotic transformations of murine vertebrae and concomitant alteration of *Hox* codes induced by retinoic acid. *Cell* **67**, 89–104 (1991).
3. Schumacher, A. & Magnuson, T. Murine *Polycomb*- and *trithorax*-group genes regulate homeotic pathways and beyond. *Trends Genet.* **13**, 167–170 (1997).
4. Abate-Shen, C. Deregulated homeobox gene expression in cancer: cause or consequence? *Nature Rev. Cancer* **2**, 777–785 (2002).
5. Wohlschlegel, J. A. *et al.* Inhibition of eukaryotic DNA replication by Geminin binding to Cdt1. *Science* **290**, 2309–2312 (2000).
6. Tada, S., Li, A., Maiorano, D., Méchal, M. & Blow, J. J. Repression of origin assembly in metaphase depends on inhibition of RLF-B/Cdt1 by Geminin. *Nature Cell Biol.* **3**, 107–113 (2001).
7. Kroll, K. L., Salic, A. N., Evans, L. M. & Kirschner, M. W. Geminin, a neutralizing molecule that demarcates the future neural plate at the onset of gastrulation. *Development* **125**, 3247–3258 (1998).
8. Bornemann, D., Miller, E. & Simon, J. The *Drosophila* Polycomb group gene *Sex comb* on midleg (*Scm*) encodes a zinc finger protein with similarity to polyhomeotic protein. *Development* **122**, 1621–1630 (1996).
9. Tomotsune, D. *et al.* A novel member of murine Polycomb-group proteins, *Sex comb* on midleg homolog protein, is highly conserved, and interacts with *Rae28/Mph1* *in vitro*. *Differentiation* **65**, 229–239 (1999).
10. Takihara, Y. *et al.* Targeted disruption of the mouse homologue of the *Drosophila* polyhomeotic gene leads to altered anteroposterior patterning and neural crest defects. *Development* **124**, 3673–3682 (1997).
11. Akasaka, T. *et al.* A role of *mel-18*, a Polycomb group-related vertebrate gene, during the anteroposterior specification of the axial skeleton. *Development* **122**, 1513–1522 (1996).
12. Kulartz, M. *et al.* Expression and phosphorylation of the replication regulator protein geminin. *Biochem. Biophys. Res. Commun.* **305**, 412–420 (2003).
13. Barna, M. *et al.* Plzf mediates transcriptional repression of *HoxD* gene expression through chromatin remodeling. *Dev. Cell* **3**, 499–510 (2002).
14. Pekarik, V. *et al.* Screening for gene function in chicken embryo using RNAi and electroporation. *Nature Biotechnol.* **21**, 93–96 (2003).
15. McGarry, T. J. & Kirschner, M. W. Geminin, an inhibitor of DNA replication, is degraded during mitosis. *Cell* **93**, 1043–1053 (1998).
16. Del Bene, F., Tessmar-Raible, K. & Wittbrodt, J. Direct interaction of geminin and Six3 in eye development. *Nature* **427**, 745–749 (2004).
17. Care, A. *et al.* HoxB7 constitutively activates basic fibroblast growth factor in melanomas. *Mol. Cell. Biol.* **16**, 4842–4851 (1996).
18. Duboule, D. Vertebrate *Hox* genes and proliferation: an alternative pathway to homeosis? *Curr. Opin. Genet. Dev.* **5**, 525–528 (1995).
19. Tetsu, O. *et al.* Mel-18 negatively regulates cell cycle progression upon B cell antigen receptor stimulation through a cascade leading to *c-myc/cdc25*. *Immunity* **9**, 439–448 (1998).
20. Krosi, J. *et al.* Cellular proliferation and transformation induced by HOXB4 and HOXB3 proteins involves cooperation with PBX1. *Oncogene* **16**, 3403–3412 (1998).
21. Thorsteinsdottir, U. *et al.* Overexpression of HOXA10 in murine hematopoietic cells perturbs both myeloid and lymphoid differentiation and leads to acute myeloid leukemia. *Mol. Cell. Biol.* **17**, 495–505 (1997).
22. Wohlschlegel, J. A., Kutok, J. L., Weng, A. P. & Dutta, A. Expression of geminin as a marker of cell proliferation in normal tissues and malignancies. *Am. J. Pathol.* **161**, 267–273 (2002).

23. Gabellini, D. *et al.* Early mitotic degradation of the homeoprotein HOXC10 is potentially linked to cell cycle progression. *EMBO J.* **22**, 3715–3724 (2003).
24. Suzuki, M. *et al.* Involvement of the Polycomb-group gene *Ring1B* in the specification of the anterior–posterior axis in mice. *Development* **129**, 4171–4183 (2002).
25. Saurin, A. J. *et al.* The human Polycomb group complex associates with pericentromeric heterochromatin to form a novel nuclear domain. *J. Cell Biol.* **142**, 887–898 (1998).
26. Megason, S. G. & McMahon, A. P. A mitogen gradient of dorsal midline Wnts organizes growth in the CNS. *Development* **129**, 2087–2098 (2002).
27. Shen, W.-F. *et al.* AbdB-like Hox proteins stabilize DNA binding by the Meis1 homeodomain proteins. *Mol. Cell. Biol.* **17**, 6448–6458 (1997).
28. Semenza, G. L., Wang, G. L. & Kundu, R. DNA binding and transcriptional properties of wild-type and mutant forms of the homeodomain protein Mx2. *Biochem. Biophys. Res. Commun.* **209**, 257–262 (1995).
29. Elbashir, S. M., Harborth, J., Weber, K. & Tuschl, T. Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods* **26**, 199–213 (2002).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank D. Duboule, A. Dutta, C. J. Tabin, S. Potter, R. Maas and H. Koseki for plasmids and antibodies, M. Pitulescu for discussions, J. Wittbrodt for communication and discussion of unpublished results, and P. Gruss and D. Gallwitz for their continuous support. This work was funded by the Max Planck Society and a DFG grant.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.K. (mkessel1@gwdg.de).

An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay

Isabel M. Palacios¹, David Gatfield², Daniel St Johnston¹ & Elisa Izaurralde²

¹Wellcome Trust/Cancer Research UK Institute and Department of Genetics, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK
²EMBL, Meyerhofstrasse 1, D-69117 Heidelberg, Germany

The specification of both the germ line and abdomen in *Drosophila* depends on the localization of *oskar* messenger RNA to the posterior of the oocyte^{1,2}. This localization requires several *trans*-acting factors, including Barentsz and the Mago–Y14 heterodimer, which assemble with *oskar* mRNA into ribonucleoprotein particles (RNPs) and localize with it at the posterior pole^{3–7}. Although Barentsz localization in the germ line depends on Mago–Y14, no direct interaction between these proteins has been detected⁵. Here, we demonstrate that the translation initiation factor eIF4AIII interacts with Barentsz and is a component of the *oskar* messenger RNP localization complex. Moreover, eIF4AIII interacts with Mago–Y14 and thus provides a molecular link between Barentsz and the heterodimer. The mammalian Mago (also known as Magoh)–Y14 heterodimer is a component of the exon junction complex^{8–11}. The exon junction complex is deposited on spliced mRNAs and functions in non-sense-mediated mRNA decay (NMD)^{9,11–14}, a surveillance mechanism that degrades mRNAs with premature translation-termination codons. We show that both Barentsz and eIF4AIII are essential for NMD in human cells. Thus, we have identified eIF4AIII and Barentsz as components of a conserved protein complex that is essential for mRNA localization in flies and NMD in mammals.

To understand further how Barentsz (Btz) is recruited to *oskar* mRNPs, we screened a *Drosophila* two-hybrid complementary DNA library¹⁵ with the amino-terminal conserved 399 amino acids of Btz as a bait. This screen identified a protein sharing 82% identity with

Xenopus eIF4AIII (refs 16, 17). *Drosophila* eIF4AIII is 66% identical to *Drosophila* eIF4AI and contains all the conserved motifs of the DEAD box protein family (Supplementary Fig. 1a)¹⁸. Despite this conservation, *Drosophila* eIF4AI does not interact with Btz (Fig. 1a).

A search of potential *eIF4AIII* mutant alleles identified the lethal complementation group *l(3)84Fi* (ref. 19) as an *eIF4AIII* allele, which we call *eIF4AIII*¹⁹, with a glycine to serine substitution at position 150 (Supplementary Fig. 1a, b). This glycine is conserved in all the eIF4AI and eIF4AIII proteins, and lies immediately adjacent to motif Ib, which is essential for helicase activity (Supplementary Fig. 1a). To investigate the role of *eIF4AIII* in the localization of *oskar* mRNA, we generated *eIF4AIII*¹⁹ mutant clones in the germ line. However, we were unable to obtain *eIF4AIII*¹⁹ homozygous oocytes, and only mosaic egg chambers containing both mutant and wild-type cells survived (Fig. 2a, b). In these cysts, the mutant cells show an aberrant organization of the actin cytoskeleton, nuclei of reduced size, and a detachment of the follicle cells, which is indicative of cell death (Fig. 2a, b).

In agreement with the interaction observed in the two-hybrid system, Btz and eIF4AIII show a strong genetic interaction. The

phenotype of homozygotes for the hypomorphic allele *btz*¹ is strongly enhanced when these flies are also heterozygous for *eIF4AIII*¹⁹. For instance, *btz*¹ egg chambers show a partially penetrant defect in *oskar* mRNA localization⁵ (Fig. 2, compare panels e and d) but the amount of localized mRNA is markedly reduced when the females are also heterozygous for *eIF4AIII*¹⁹ (Fig. 2f). Similar results were obtained by analysing the distribution of Staufen (Stau) protein, which co-localizes with *oskar* mRNA²⁰ (Supplementary Fig. 2a–f). As the posterior localization of *oskar* mRNA defines where the germline determinants are localized^{1,2}, most embryos from *btz*¹ females have fewer pole cells than wild type, and this leads to loss of one or both ovaries in 6% of the female progeny (Fig. 2g). The progeny of mothers homozygous for *btz*¹ and heterozygous for *eIF4AIII*¹⁹ show a much stronger phenotype, in which 31% lack at least one ovary (Fig. 2g). Finally, embryos produced by homozygous *btz*¹ females with one mutant copy of *oskar* (for example, *osk*⁸⁴ allele) show a partial loss of abdominal segments, and this phenotype is strongly enhanced by the presence

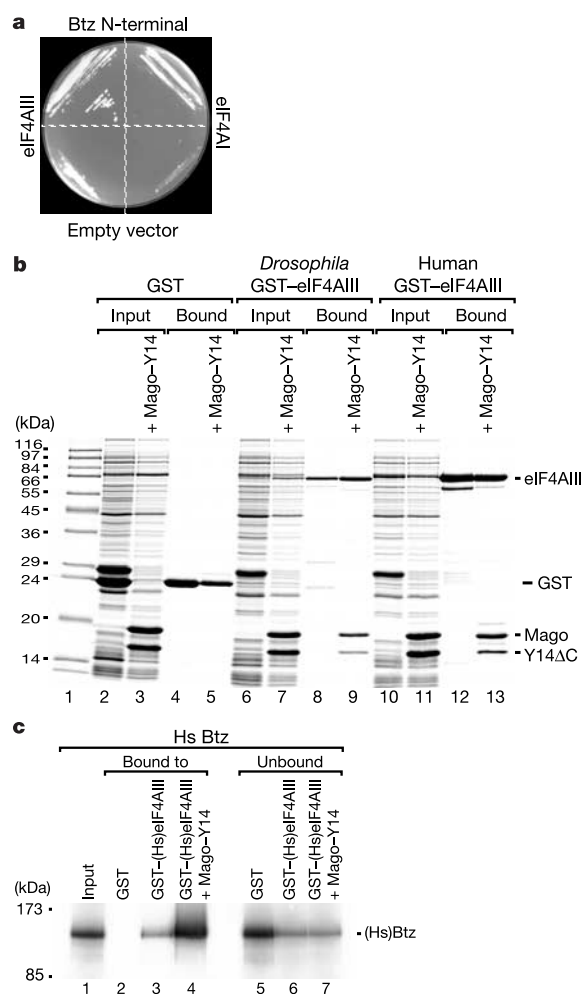


Figure 1 Identification of eIF4AIII as a partner of Btz and Mago-Y14. **a**, eIF4AIII, but not eIF4AI, interacts with Btz in a yeast two-hybrid assay. The growth of cells near the centre of the plate is an indicator of interaction. **b**, GST pull-down assays with lysates from *E. coli* expressing the proteins indicated above the lanes. One-fiftieth of the inputs and one-third of the bound fractions were analysed by SDS–polyacrylamide gel electrophoresis (PAGE) followed by Coomassie blue staining. **c**, GST pull-down assays with [³⁵S]methionine-labelled human Btz and the recombinant proteins indicated above the lanes. One-tenth of the input or of the supernatants and one-third of the bound fractions were analysed by SDS–PAGE followed by fluorography. Dm, *Drosophila melanogaster*; Hs, *Homo sapiens*.

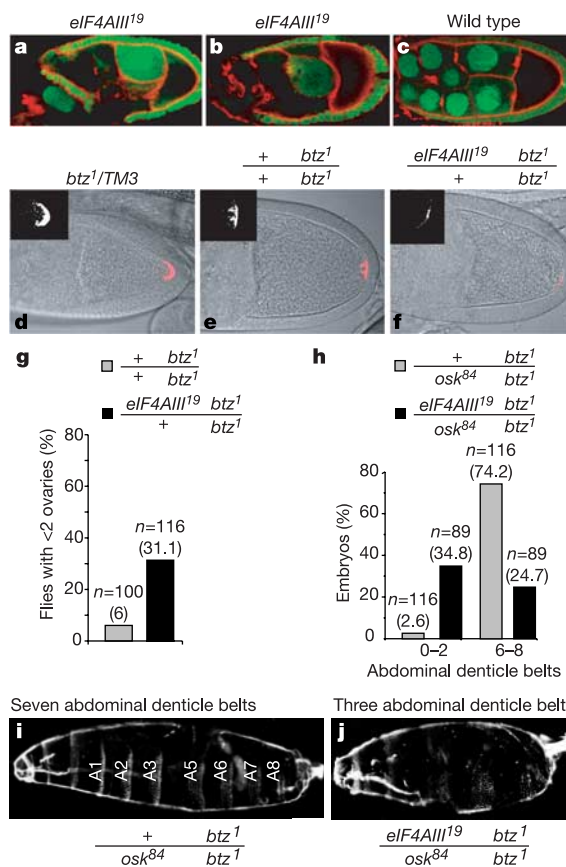


Figure 2 eIF4AIII functions with Btz in the localization of *oskar* mRNA. **a–c**, Staining of the actin cytoskeleton by rhodamine-labelled phalloidin (red) of stage 9 wild-type egg chambers (**c**), or in egg chambers where *eIF4AIII*¹⁹ germline clones were generated using the lack of GFP as a marker for homozygous mutant cells (**a**, **b**). **d–f**, Detection of *oskar* mRNA by *in situ* hybridization in *btz*¹ mutant ovaries (stage 9) with (**f**) or without (**e**) a copy of the *eIF4AIII*¹⁹ allele. *btz*¹ heterozygous flies are used as a control (**d**). Oocytes are shown in phase contrast with the fluorescent RNA probe overlaid. The insets show the probe alone. **g**, Quantification of the number of flies with less than two ovaries in the progeny of mothers homozygous for *btz*¹ with or without a copy of the *eIF4AIII*¹⁹ allele. **h**, Quantification of the number of embryos with 0–2 or 6–8 abdominal denticle belts produced by *osk*⁸⁴ *btz*¹/*btz*¹ or *osk*⁸⁴ *btz*¹/*eIF4AIII*¹⁹ *btz*¹ mothers. **i**, **j**, Representative example of abdominal denticle belts in embryos from *osk*⁸⁴ *btz*¹/*btz*¹ (**i**) and *osk*⁸⁴ *btz*¹/*eIF4AIII*¹⁹ *btz*¹ (**j**) mothers. Wild-type embryos have eight abdominal denticle belts (A1–A8).

of one mutant copy of *eIF4AIII* (Fig. 2h–j). The strength of these genetic interactions is comparable to that observed with other posterior group mutants that disrupt *oskar* mRNA localization, such as *stau* and *mago*²¹.

To study further *eIF4AIII* during oogenesis, we performed immunostaining with a specific *eIF4AIII* antibody. In the germ line, at stages 1–9 of oogenesis, the localization of *eIF4AIII* in the oocyte cytoplasm is identical to that observed for Btz, Mago–Y14, Stau and *oskar* mRNA^{1–7,20}. This is best seen in stages 1–6, in which *eIF4AIII* concentrates in a posterior crescent in the oocyte (Fig. 3a, arrowhead). At stage 9, at the time of *oskar* mRNP localization, *eIF4AIII* is found at the posterior pole of the oocyte (Fig. 3b, c, arrow). Labelling of green fluorescent protein (GFP)-tagged Mago-expressing oocytes with *eIF4AIII* antibodies revealed that the two proteins co-localize to the posterior of the oocyte (Fig. 3c). The distribution of endogenous *eIF4AIII* was confirmed in transformants that express GFP–*eIF4AIII* (Supplementary Fig. 3). The localization pattern of *eIF4AIII* within the oocyte cytoplasm at stages 1–9 of oogenesis strongly suggests that *eIF4AIII* is a component of the *oskar* mRNP localization complex.

In addition to its localization in the oocyte cytoplasm, *eIF4AIII* is detected within the nucleus of all cell types of the egg chamber (Fig. 3c, d) in a pattern that resembles the distribution of Mago–Y14 (Fig. 3d). The subcellular localization of Mago–Y14, *eIF4AIII* and Btz is conserved in mammalian cells. Indeed, whereas Btz is cytoplasmic, human Mago–Y14 localizes in the nucleoplasm and

in nuclear speckles (Fig. 3f; see also Supplementary Fig. 3 and refs 8, 10, 11, 22). As with *Drosophila* *eIF4AIII*, human *eIF4AIII* is detected in the nucleoplasm and in nuclear speckles (Fig. 3e; see also Supplementary Fig. 3).

The nuclear localization of *eIF4AIII* suggests that, together with Mago–Y14, this protein associates with *oskar* mRNA in the nucleus, providing a link between Mago–Y14 and Btz. To test this hypothesis, we investigated whether *eIF4AIII* interacts with Mago–Y14 *in vitro*. We found that *Drosophila* Mago–Y14 dimers interact with both *Drosophila* and human *eIF4AIII* (Fig. 1b), a result that provides direct evidence for a specific physical interaction. Moreover, human Btz (also known as MLN51 (ref. 22)) specifically interacts with human *eIF4AIII* (Fig. 1c) and with *eIF4AIII*–Mago–Y14 trimeric

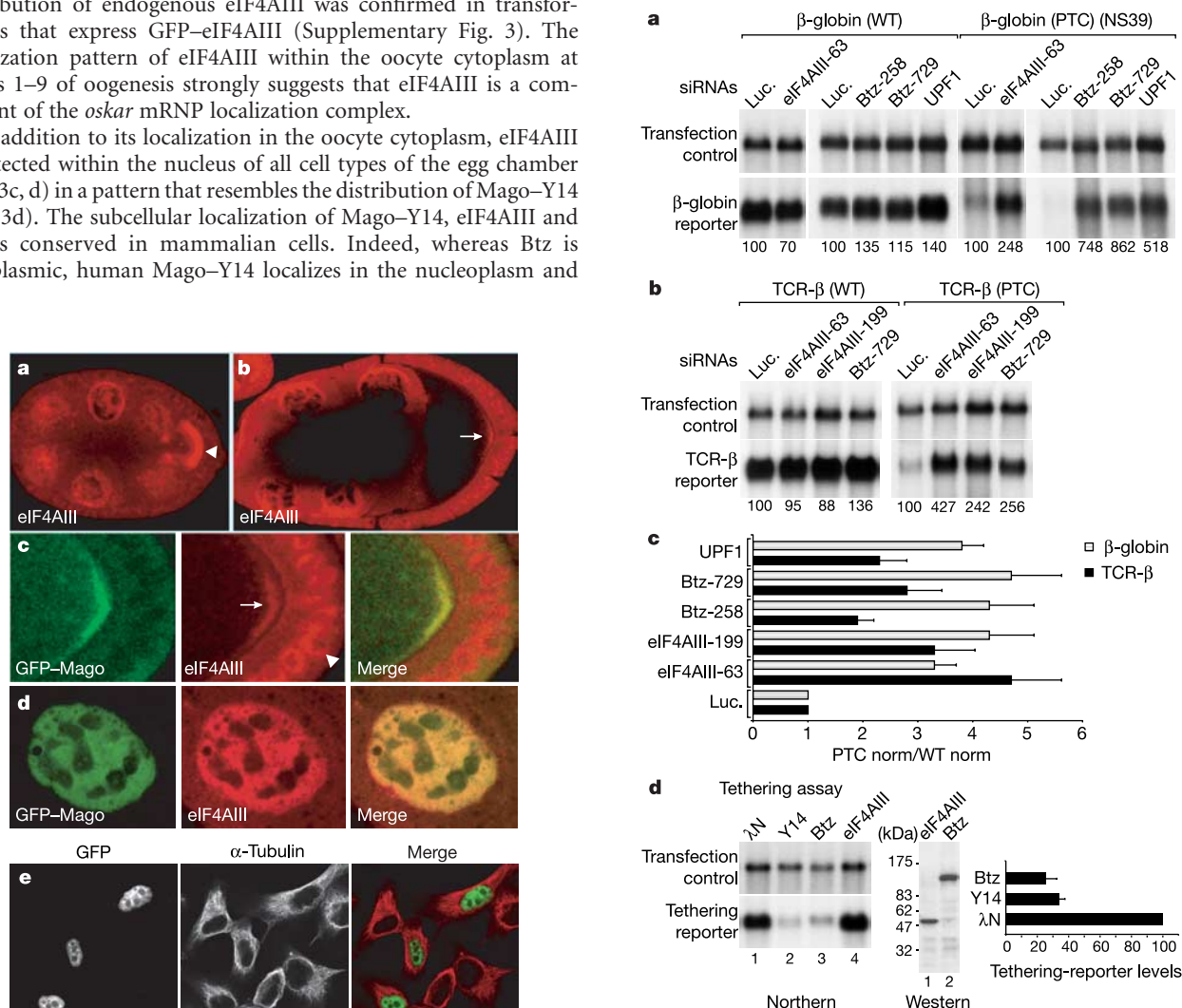


Figure 3 *eIF4AIII* localizes to the nucleus and to the posterior pole of the oocyte. **a–d**, Antibody staining for *eIF4AIII* during oogenesis. **a**, Stage 5; **b**, stage 9, **c**, posterior region of a stage 9 egg chamber expressing GFP–Mago; **d**, nurse cell nucleus of a GFP–Mago egg chamber. The arrowhead in **c** points to a follicle cell nucleus. **e, f**, HeLa cells transiently expressing GFP–*eIF4AIII* or Btz–GFP, as indicated, were stained with an anti- α -tubulin antibody (red).

Figure 4 Human *eIF4AIII* and Btz are required for NMD. **a, b**, HeLa cells were transfected with the indicated siRNAs and either the β -globin (**a**) or TCR- β (**b**) reporters. Numbers below the lanes correspond to the levels of the reporters normalized to those of the transfection control, and set to 100 in cells transfected with luciferase siRNA. PTC, premature stop codon. **c**, The normalized values (norm) obtained for the premature-stop-codon-containing mRNAs were divided by those obtained for the corresponding wild-type transcripts to correct for nonspecific effects of the depletions. The mean values obtained in at least three independent experiments $\pm$ standard deviations are shown. **d**, The levels of the tethering-NMD reporter in the presence of λ N fusions of Btz, Y14 or *eIF4AIII* were normalized to those of the transfection control and expressed as a percentage of the levels observed when the λ N peptide alone was co-expressed. The mean values $\pm$ standard deviations are shown. The expression levels of the proteins were analysed by western blot with an anti- λ N antibody¹².

complexes. These data, together with the two-hybrid interaction described above, indicate that this network of protein interactions is conserved.

Structural and functional studies indicate that human Mago-Y14 is required for NMD^{12,13}. The conservation of the physical interactions between Mago-Y14, Btz and eIF4AIII, and of their subcellular localization, prompted us to investigate whether human eIF4AIII and Btz function in NMD. We therefore analysed the steady-state levels of a previously described β -globin mRNA reporter (NS39)²³ carrying a premature stop codon, in HeLa cells depleted of Btz or eIF4AIII by RNA interference (RNAi). In cells treated with a short interfering (si)RNA directed against an irrelevant target (luciferase), the levels of the NS39 transcript were reduced in abundance to approximately 5% of the wild-type β -globin levels due to its degradation by the NMD machinery, as reported previously²³ (Fig. 4a). In contrast, depletion of eIF4AIII or Btz by RNAi increased the levels of the mRNA containing the premature stop codon (Fig. 4a), indicating that NMD is impaired. The upregulation of NS39 mRNA ranged between 2–5-fold in all cases (Fig. 4c). Similar values were obtained when a different reporter, derived from the T-cell receptor- β (TCR- β) minigene²⁴, was analysed (Fig. 4b, c). These values are in agreement with those obtained for this and other NMD reporters in human cells depleted of bona fide NMD factors such as UPF1 or Y14 (Fig. 4a, c; see also refs 12, 24, 25). These results provide compelling evidence for a role of human eIF4AIII and Btz in NMD. Of note, depletion of Mago-Y14, eIF4AIII or Btz from *Drosophila* Schneider cells does not lead to the stabilization of mRNAs containing a premature stop codon (data not shown), as exon-exon boundaries are not used to define premature stop codons in this organism²⁵.

The role of Btz in NMD was confirmed further in an NMD-tethering assay. In this assay, proteins fused to the λ N peptide are co-transfected with a β -globin-derived reporter harbouring five high-affinity λ N-peptide-binding sites downstream of the stop codon¹². Degradation of this reporter is elicited by tethering NMD factors to these sites^{12,13}. Tethering Btz to the reporter results in its degradation to a similar extent as observed for Y14 or Mago (Fig. 4d)^{12,13}, indicating that Btz is a bona fide component of the NMD machinery. In contrast, tethering eIF4AIII does not elicit NMD (Fig. 4d). Thus, Btz, like Y14–Mago, can recruit additional components of the NMD machinery, thereby promoting degradation of the reporter. Although we cannot rule out that eIF4AIII fused to the λ N peptide is inactive, it is also possible that this protein is required for assembly of the exon junction complex, but not for the recruitment of the NMD machinery in the cytoplasm.

Here we have identified *Drosophila* eIF4AIII as a functional component of the *oskar* mRNP localization complex. At steady state, eIF4AIII and Mago-Y14 are nuclear whereas Btz is cytoplasmic, suggesting that the assembly of *oskar* mRNPs occurs in a stepwise manner with Mago-Y14 and eIF4AIII binding *oskar* mRNA in the nucleus and Btz being recruited via eIF4AIII as the complex is exported into the cytoplasm. Several lines of evidence support this model. First, eIF4AIII interacts with both Mago-Y14 and Btz. Second, in oocytes the localization of eIF4AIII precisely overlaps that of Mago-Y14, Btz, Stau and *oskar* mRNA. Finally, the eIF4AIII¹⁹ mutant allele enhances the *oskar* mRNA localization phenotype of a *btz* hypomorph allele. As Mago-Y14 and eIF4AIII are components of the exon junction complex in mammals^{8–11,26} (M. Moore and N. Sonenberg, personal communication), and the splicing of *oskar* mRNA is required for its localization (O. Hachet and A. Ephrussi, personal communication), it seems likely that Mago-Y14 and eIF4AIII are components of the exon junction complex in *Drosophila*, and that they are recruited onto *oskar* mRNA as it is spliced in the nucleus.

Recently, murine Btz has been shown to localize in dendrites together with RNP granules containing Stau²². These observations suggest that the role of Btz in mRNA localization is conserved, and it

would be interesting to determine whether mammalian eIF4AIII and Mago-Y14 are also involved in mRNA localization. Our findings reveal that a conserved protein complex comprising Mago-Y14, eIF4AIII and Btz has multiple roles in post-transcriptional mRNA metabolism. Notably, the complex is essential for mRNA localization in *Drosophila* and for mRNA surveillance in mammalian cells. □

Methods

Two-hybrid screen and protein interactions

The two-hybrid screen was performed using *Saccharomyces cerevisiae* strain L40 (ref. 27) and an oligo(dT)-primed *Drosophila* ovarian cDNA library¹⁵. The LexA–Btz bait comprises the first 399 amino acids of Btz. A plasmid allowing the expression of untagged *Drosophila* Mago–Y14 Δ C in *Escherichia coli* has been described previously¹³. Plasmids allowing the expression of glutathione S-transferase (GST) fusions of *Drosophila* or human eIF4AIII were generated by inserting the corresponding cDNAs at the *NcoI*/*Bam*HI or *XhoI*/*NotI* sites of vectors pGEXCS or pGEX5X-1, respectively. Proteins were expressed in the *E. coli* BL21 strain. GST pull-down assays were performed as described previously^{8,13}.

Fly strains and genetics

All germline clones were generated by the FRT/FLP recombinase system²⁸, using either the lack of GFP as a marker for homozygous clones or the ovoD system, in which only homozygous mutant cells develop further than stage 4/5 of oogenesis. Germline clones were generated by heat-shocking third instar larvae for 2 h at 37 °C for three consecutive days. The generation of the eIF4AIII¹⁹ mutant allele is described in Supplementary Fig. 1. We used the following fly stocks: *btz*¹/TM3, eIF4AIII¹⁹/TM3, FRT82B eIF4AIII¹⁹/TM3, eIF4AIII¹⁹ *btz*¹/TM3, *osk*²⁴ *btz*¹/TM3, Btz–GFP, GFP–Mago² and GFP–eIF4AIII.

Molecular cloning

The *Drosophila* eIF4AIII coding region was amplified and cloned into pD277–GFP6 (ref. 5) to generate a transformation construct in which the α 4-tubulin promoter drives germline-specific expression of GFP fused to eIF4AIII.

Immunocytochemistry and *in situ* hybridization

Antibodies were raised in rabbits against an N-terminal eIF4AIII peptide (NMARKNAQAEDLSNVE) specific for *Drosophila* eIF4AIII. The purified antibody recognizes a unique band of the right size in western blots of ovarian extracts (data not shown). Antibody stainings and RNA *in situ* hybridizations were performed as previously reported⁵. In Fig. 2d–f, the *oskar* mRNA antisense probe was *in-vitro*-transcribed in the presence of aminoallyl-UTP (Sigma), and was directly labelled with the Amersham Cy3 Mono-reactive Dye Pack. For expression in HeLa cells as GFP fusions, human Y14, Mago and eIF4AIII cDNAs were cloned in pEGFP-C1 vector (Clontech). A plasmid expressing Btz–GFP is described in ref. 22. Cells were fixed and stained with anti- α -tubulin monoclonal antibody (clone DM 1A, Sigma) and a Texas-red-coupled secondary antibody, 16 h after transfection.

RNAi and NMD assays in HeLa cells

Depletion of endogenous proteins by RNAi, western blot analysis and the NMD assays were carried out as described previously^{13,25}. The specificity and efficiency of the siRNAs was tested by western blot as shown in Supplementary Fig. 4. The sequences targeted by the siRNAs are given in Supplementary Fig. 4.

Received 14 November 2003; accepted 19 January 2004; doi:10.1038/nature02351.

- Ephrussi, A., Dickinson, L. K. & Lehmann, R. *Oskar* organizes the germ plasm and directs localisation of the posterior determinant *nanos*. *Cell* **66**, 37–50 (1991).
- Kim-Ha, J., Smith, J. L. & Macdonald, P. M. *oskar* mRNA is localized to the posterior pole of the *Drosophila* oocyte. *Cell* **66**, 23–35 (1991).
- Micklem, D. R. *et al.* The *mago nashi* gene is required for the polarisation of the oocyte and the formation of perpendicular axes in *Drosophila*. *Curr. Biol.* **7**, 468–478 (1997).
- Newmark, P. A., Mohr, S. E., Gong, L. & Boswell, R. E. *mago nashi* mediates the posterior follicle cell-to-oocyte signal to organize axis formation in *Drosophila*. *Development* **124**, 3197–3207 (1997).
- van Eeden, F. J., Palacios, I. M., Petronczki, M., Weston, M. J. & St Johnston, D. Barentsz is essential for the posterior localisation of *oskar* mRNA and colocalizes with it to the posterior pole. *J. Cell Biol.* **154**, 511–523 (2001).
- Mohr, S. E., Dillon, S. T. & Boswell, R. E. The RNA-binding protein Tsunagi interacts with Mago Nashi to establish polarity and localize *oskar* mRNA during *Drosophila* oogenesis. *Genes Dev.* **15**, 2886–2899 (2001).
- Hachet, O. & Ephrussi, A. *Drosophila* Y14 shuttles to the posterior of the oocyte and is required for *oskar* mRNA transport. *Curr. Biol.* **11**, 1666–1674 (2001).
- Le Hir, H., Gatfield, D., Braun, I. C., Forler, D. & Izaurralde, E. The protein Mago provides a link between splicing and mRNA localisation. *EMBO Rep.* **2**, 1119–1124 (2001).
- Le Hir, H., Gatfield, D., Izaurralde, E. & Moore, M. J. The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J.* **20**, 4987–4997 (2001).
- Kataoka, N., Diem, M. D., Kim, V. N., Yong, J. & Dreyfuss, G. Magoh, a human homolog of *Drosophila* Mago Nashi protein, is a component of the splicing-dependent exon-exon junction complex. *EMBO J.* **20**, 6424–6433 (2001).
- Kim, V. N. *et al.* The Y14 protein communicates to the cytoplasm the position of exon-exon junctions. *EMBO J.* **20**, 2062–2068 (2001).

12. Gehring, N. H., Neu-Yilik, G., Schell, T., Hentze, M. W. & Kulozik, A. E. Y14 and hUpf3b form an NMD-activating complex. *Mol. Cell* **11**, 939–949 (2003).
13. Fribourg, S., Gatfield, D., Izaurralde, E. & Conti, E. A novel mode of RBD-protein recognition in the Y14-Mago complex. *Nature Struct. Biol.* **10**, 433–439 (2003).
14. Lykke-Andersen, J., Shu, M. D. & Steitz, J. A. Communication of the position of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1. *Science* **293**, 1836–1839 (2001).
15. Grosshans, J., Schnorrer, F. & Nusslein-Volhard, C. Oligomerisation of Tube and Pelle leads to nuclear localisation of dorsal. *Mech. Dev.* **81**, 127–138 (1999).
16. Weinstein, D. C., Honore, E. & Hemmati-Brivanlou, A. Epidermal induction and inhibition of neural fate by translation initiation factor 4AIII. *Development* **124**, 4235–4242 (1997).
17. Li, Q. *et al.* Eukaryotic translation initiation factor 4AIII (eIF4AIII) is functionally distinct from eIF4AI and eIF4AII. *Mol. Cell. Biol.* **19**, 7336–7346 (1999).
18. Tanner, N. K., Cordin, O., Banroques, J., Doere, M. & Linder, P. The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. *Mol. Cell* **11**, 127–138 (2003).
19. Baker, B. S., Hoff, G., Kaufman, T. C., Wolfner, M. F. & Hazelrigg, T. The *doublesex* locus of *Drosophila melanogaster* and its flanking regions: a cytogenetic analysis. *Genetics* **127**, 125–138 (1991).
20. St Johnston, D., Beuchle, D. & Nusslein-Volhard, C. Staufén, a gene required to localize maternal RNAs in the *Drosophila* egg. *Cell* **66**, 51–63 (1991).
21. Smith, R. *Screens to Find Novel Genes Involved in Pole Plasm Formation in Drosophila melanogaster*. PhD Thesis, Cambridge Univ. (1998).
22. Macchi, P. *et al.* Barentsz, a new component of the Staufén-containing ribonucleoprotein particles in mammalian cells, interacts with Staufén in an RNA-dependent manner. *J. Neurosci.* **23**, 5778–5788 (2003).
23. Thermann, R. *et al.* Binary specification of nonsense codons by splicing and cytoplasmic translation. *EMBO J.* **17**, 3484–3494 (1998).
24. Mendell, J. T., ap Rhys, C. M. & Dietz, H. C. Separable roles for rent1/hUpf1 in altered splicing and decay of nonsense transcripts. *Science* **298**, 419–422 (2002).
25. Gatfield, D., Unterholzner, L., Ciccarelli, F. D., Bork, P. & Izaurralde, E. Nonsense-mediated mRNA decay in *Drosophila*: at the intersection of the yeast and mammalian pathways. *EMBO J.* **22**, 3960–3970 (2003).
26. Chan, C. C. *et al.* eIF4A3 is a novel component of the exon junction complex. *RNA* **10**, 200–209 (2004).
27. Hollenberg, S. M., Sternglanz, R., Cheng, P. F. & Weintraub, H. Identification of a new family of tissue-specific basic helix-loop-helix proteins with a two-hybrid system. *Mol. Cell. Biol.* **15**, 3813–3822 (1995).
28. Chou, T. B. & Perrimon, N. The autosomal FLP-DFS technique for generating germline mosaics in *Drosophila melanogaster*. *Genetics* **144**, 1673–1679 (1996).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We acknowledge G. Dreyfuss, A. Ephrussi, M. Moore and N. Sonenberg for sharing unpublished observations. We thank D. Micklem for the gift of GFP–Mago flies; M. Hentze, A. Kulozik and M. Kiebler for plasmids and antibodies; R. Cantera for help in collecting confocal images; and P. Lawrence for critical reading of the manuscript. I.M.P. was supported by the Royal Society Dorothy Hodgkin Fellowship. D.S.J. was supported by a Wellcome Trust Principal Research Fellowship. D.G. and E.I. are supported by the European Molecular Biology Organization.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to D.S.J. (ds139@mole.bio.cam.ac.uk) or E.I. (izaurralde@embl-heidelberg.de).

Contacts

Publisher: Ben Crowe
Editor: Paul Smaglik
Marketing Manager: David Bowen

European Head Office, London

The Macmillan Building
4 Crinan Street
London N1 9XW, UK
Tel +44 (0) 20 7843 4961
Fax +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

Naturejobs Sales Director:
Nevin Bayoumi (4978)

UK/ RoW/ Ireland:

Matt Powell (4953)
Andy Douglas (4975)
Frank Phelan (4944)

Scandinavia/ Spain/ Portugal:

Evelina Rubio Håkansson (4973)

Natureevents: Silke Opstrup (4994)

France/ Switzerland:

Amelie Pequinot (4974)

Production Manager:

Billie Franklin
To send materials use London
address above.
Tel +44 (0) 20 7843 4814
Fax +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

Naturejobs web development:

Tom Hancock

Naturejobs online production:

Ben Lund

European Satellite Office

Germany/ Austria/ Netherlands/
Italy/ Belgium:

Patrick Phelan, Odo Wulffen
Tel + 49 89 54 90 57 11/-2
Fax + 49 89 54 90 57 20
e-mail: p.phelan@nature.com
o.wulffen@nature.com

US Head Office, New York

345 Park Avenue South,
10th Floor, New York, NY 10010-1707
Tel +1 800 989 7718
Fax +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager:

Peter Bless

US Advertising Coordinator:

Linda Adam

Japan Head Office, Tokyo

MG Ichigaya Building (5F),
19-1 Haraikatomachi,
Shinjuku-ku,
Tokyo 162-0841
Tel +81 3 3267 8751
Fax +81 3 3267 8746

Asia-Pacific Sales Director:

Rinoko Asami
e-mail: rasami@naturejpn.com

naturejobs

Pitching ideas

In the United States, sandlots are places where aspiring athletes toss baseballs. But next autumn, the idea will cross the Atlantic to the University of Bristol, UK, where sandlots will become an arena for students to pitch their ideas about translational research. The course will see postgraduates from different disciplinary backgrounds team up in an attempt to solve a series of problems centred on the gap between basic research and practical healthcare. The results will be presented to a panel of international academics, industrial researchers and scientists from the UK National Health Service for discussion and comment.

Beverley Randle, director of Bristol's Healthcare Enterprise Programme, got the idea for the course while participating in a similar 'sandpit' initiative last year (see *Nature* 427, 187; 2004). She was intrigued by the way researchers from maths, physics, chemistry and biology seemed to invent a "scientific Esperanto" to communicate across disciplines. She wondered whether "throwing students into the deep end" of problem solving would be a good way to educate scientists for an interdisciplinary world.

She hopes that the new programme will solve a few other problems, as well. First, there's a shortage of people poised to do 'bench-to-bedside' research because current systems encourage scientists to choose between patient care and clinical research. The new scheme offers an alternative, by letting students pick which parts of the clinical-research system they want to be involved with early on. And the curriculum, which makes students come up with a final project, ranging from a business plan to a research paper, addresses the issue of how to make the course relevant to the students' future. "We want to tailor this to the students' needs so they can carry this into their careers," says Randle.

The contest should be interesting to watch from the sidelines. And if it's successful, expect more universities to give the game a go.

Paul Smaglik
Naturejobs editor



Contents

POSTDOCS AND STUDENTS

The pros and cons of
short-term contracts p760

CAREER VIEW

Scientists & Societies
Improving the lot of US
postdocs
Graduate Journal
Home or abroad?
Movers
Paul Billings p762

WWW.NATUREJOBS.COM

Career centre
Information on the
scientific job market

FOCUS

SPOTLIGHT

RECRUITMENT

ANNOUNCEMENTS

EVENTS

Short-term limbo

In Europe, short-term contracts are no longer just for the first postdoc. Karen Kreeger looks at the bewildering variety of terms and conditions on offer.

Short-term contracts for young scientists, as opposed to the tenure-track model with a permanent position at the end, are becoming more popular as an employment option for some institutions. But they set up an inevitable conflict between what's good for science and the research institution — productivity and new sources of creativity — and what's good for the scientist — job security.

Especially in Europe, where these arrangements are most popular, young scientists must weigh the pros and cons before committing themselves. As a complication, although many institutions offer fixed-term contracts, actual terms and conditions vary widely, sometimes blurring the distinction between postdocs and junior faculty members.

For example, the Friedrich Miescher Institute for Biomedical Research in Basel, Switzerland, hires junior group leaders on a three-year contract, renewable for another three years, with the possibility of a tenure-track position only at the end. Meanwhile, at the Institute of Molecular Pathology (IMP) in Vienna, Austria, investigators start with a five-year contract, which can be renewed to eight years depending on the outcome of an internal review, but these contracts rarely lead to a tenured position. And the Wellcome Trust/Cancer Research UK Gurdon Institute in Cambridge, offers two levels of appointment — group leaders have a five-year initial appointment, renewable for another five years after which they either get a separate senior position or have to leave; whereas senior group leaders are, in effect, appointed on a permanent basis as long as their programme's grant is renewed at the five-yearly review.

Depending on your perspective, short-term contracts can have conflicting benefits for the

institution, the employee and perhaps even for scientific research. For an institution, the advantage is that it is “renewing itself almost constantly”, says Jan-Michael Peters, a senior scientist at the IMP and an early exception to the rule that prevented short-term junior faculty members at the IMP from securing long-term senior positions there — he and a colleague were the first to jump from junior to senior positions two years ago. “The average age of the faculty is relatively young, which helps to make it a lively, energetic and fun place,” he says.

But, he counters, the cons are also obvious. “If someone is a really good colleague and an outstanding scientist it is sad to let him or her go,” he says. At the IMP this has been solved by rare promotions or by placing colleagues in senior positions at neighbouring institutions. “But there are still excellent colleagues who we have to let go, simply because the IMP is too small to accommodate every good group leader as a senior,” he explains.

Other advantages, notes Bernd-Uwe Jahn, administrative director for the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany, is that an organization retains flexibility in adapting research programmes, bringing in fresh ideas, and taking expertise back to the scientific community through mobility. But the continual turnover does lead to loss of expertise, especially in service areas such as administration and technical staff. EMBL is an extreme example: 90% of its staff are on fixed-term contracts (see *Nature* **423**, 364–365; 2003).

Although these contracts allow for flexibility in hiring, they may backfire if they keep the best scientists away, says Jonathon Pines, senior group leader at the Gurdon Institute. Pines sees the career structure of short-term contracts as increasingly unattractive to European scientists, especially compared with US institutions. This could translate into difficulty in recruiting and retaining the right people.

MIXED MESSAGES FOR EMPLOYEES

For junior group leaders and their staff, a system of generous support at the outset in conjunction with short-term contracts provides great opportunities to pursue important scientific questions that may require expensive approaches, says Antoine Peters, a staff scientist in Thomas Jenuwein's group at the IMP. Peters started his second postdoc there in 2000 and in May he will leave to set up his own research group at the Friedrich Miescher Institute.

But, he adds, short-term contracts stop people from establishing much of a social life outside work, because the uncertainty of their future tends to drive them to work nights and weekends to get the publications that could land them a more secure position.

“The problem here is the length of contract for these short jobs,” says Tony Hyman, director of the Max Planck Institute of Molecular Cell Biology and

Thomas Jenuwein
(holding the ‘T’)
feels that short-term
contracts can hamper
a postdoc's social life.





Beat the clock: Harold Lloyd tries to tame time in *Safety Last!* Young postdocs on fixed-term contracts must also scramble to secure a steadier position.

GETTY IMAGES

Genetics in Dresden, Germany. "In Germany they have been only five years with no renewal." He adds that this is too short a time in which to set up an ambitious research programme, as an investigator who has to start looking for a job after four years. Pines agrees: "You can be out of a job at any time in your career. It's difficult to pursue ambitious projects without obvious short-term results." But on the other hand, the continual peer review keeps standards high and is a valuable exercise for the principal investigator in assessing the state and overall direction of research.

Jan-Michael Peters says that limited-term contracts do add pressure to perform into the daily mix. "This can be tough, but is not necessarily a bad thing because it keeps us all on our toes," he says. But as well as no security, these temporary positions place high demands on family members, "because moving around, changing schools and jobs can be tough for spouses and kids", he adds.

GOOD FOR SCIENCE

The mobility created by short-term contracts does have benefits for science. "They prevent an institution from stagnating; mobility of scientists promotes creativity," says Jahn. But, counters Pines, the lack of job security when combined with low pay makes academic science unattractive as a career compared with alternatives in industry, commerce or finance. "At present we seem to

have the worst of both worlds," says Pines. "Institutions must either enhance pay or offer tenure. We can't continue to do science on the cheap and expect to attract the best and brightest to the career."

Short-term contracts make it difficult to pursue long-term goals, especially if there are few publishable results in the short term. Ambitious projects may not get funded unless both the peer review and funding committee are far-sighted and supportive.



Antoine Peters sees short-term contracts as great opportunities to pursue specific projects.

Peter Schuster, professor of theoretical chemistry at Vienna University in Austria, sees a way out: a hybrid system with a mixture of short- and long-term positions. Schuster is in favour of renewable contracts for young scientists with an upper limit of eight years, if the arrangements meet two criteria.

"Junior professors must have their own budget, some limited personnel depending on the research area, and a low teaching load; and the job situation has to be sufficiently good to provide a high chance for all good people to receive an offer of a full professorship," he says. For the future, Vienna University foresees short-term contract positions for junior professors being roughly equal in number with full professorships, which would have a higher teaching load.

Although the academic tenure model is alive and kicking in most research institutions, a system that combines this approach with short-term contracts seems to be the future for more and more labs.

Karen Kreeger is a science writer based in Media, Pennsylvania.

Further reading

Nature Mater. **3**, 1 (2004).
Frazzetto, G. *EMBO Rep.* **4**, 1110–1112 (2003).
Dillon, N. *EMBO Rep.* **4**, 2–4 (2003).
Langenberg, H. *Nature* **410**, 849–850 (2001).

Web links

IMP, Vienna
♦ www.imp.univie.ac.at
Friedrich Miescher Institute for Biomedical Research, Basel
♦ www.fmi.ch
The Wellcome Trust/Cancer Research UK Gurdon Institute, Cambridge
♦ www.welc.cam.ac.uk
EMBL, Heidelberg
♦ www.embl-heidelberg.de
Max Planck Institute of Molecular Cell Biology and Genetics
♦ www.mpi-cbg.de

GRADUATE JOURNAL

Home or abroad?

I have always treasured the global nature of science. But now, as I try to decide between a postdoc in Paris and a similar position in the United States, this beloved characteristic is hampering my choice.

When I started at graduate school, the thought of going abroad for a fellowship never crossed my mind. But my time at Rockefeller University in New York has changed my thinking. Here, as an African-American and a US citizen, I am actually in a minority among my classmates in terms of both ethnicity and nationality. The plethora of positive interactions that I have had with the community of international scientists on campus has given me a fresh view on working overseas.

One factor that may weigh heavily in my final decision is the level and stability of government funding for scientific research in the respective countries. The recent protest by French scientists in response to their government's spending plans was alarming. Other aspects I am considering include overall quality of life, cultural surroundings and the feasibility of obtaining a faculty appointment. As the quality of science is excellent in both places, making a decision will be hard. Certainly, one good thing has emerged from my dilemma — I pay a lot more attention to how the dollar is doing versus the euro than I used to. ■

Tshaka Cunningham is a fifth-year graduate student at Rockefeller University in New York.

SCIENTISTS & SOCIETIES

Protecting the public investment

The training of young US scientists as they move through the academic system is supported by an investment of limited public funds from government and private foundations. As a US taxpayer, I want to be assured that this investment is recruiting the highest level of talent and providing the best training environment for our future scientific leaders.

As a postdoc, I see great potential for improving the return on this investment. It is the responsibility of the whole scientific community to take action to ensure that postdoc training is as good as it can be. Numerous groups and individuals have defined steps that could be taken, including increasing postdoc productivity and creativity by improving mentoring and professional

development programmes; recruiting and retaining the highest level of talent by boosting the compensation and benefit packages to meet the needs of our maturing postdoc population; and restoring the motivation of postdocs who feel uncertain about their future by improving career-development training. Unless we address such issues, we risk losing our brightest and best to other enterprises.

All levels of the scientific community must come together to improve and enhance postdoctoral training. To aid this process, the US National Postdoctoral Association (NPA) will bring representatives from all levels of the community together at its second annual meeting on 16–17 April in Washington DC. The common theme for the meeting is the responsibilities and actions — at various levels, including individual,

institutional and national agency or association — for improving postdoc training. The meeting aims to pool resources, develop collaborative strategies, and enhance the interaction of the NPA with the broader scientific community. This should help to protect the public investment and continue to move the scientific community beyond the nebulous concept of 'fair' or 'we deserve it' to one of shared responsibility.

The meeting coincides with the Convocation on Enhancing the Postdoctoral Experience, which will review the status of US postdocs. Hopefully, these meetings will mean that in years to come, the postdoctoral experience will be such that we readily retain the brightest and the best. ■

Steven Wendell is a postdoc at the University of Pittsburgh and an executive board member of the US National Postdoctoral Association.
♦ www.nationalpostdoc.org

MOVERS

Paul Billings, vice-president, Laboratory Corporation of America Holdings, North Carolina



Paul Billings earned an MD learning to treat humans and a PhD studying murine immunology. But it was insurance that really changed his career.

Billings became well known for his early work on the discriminatory uses of genetic information. But in 1997, when he found himself at the helm of the VA Heart of Texas Health Care Network, he was forced to deal with "how we valued

things, how we allocated funds, who we covered and for what".

Healthcare providers were "bombarded" by new research information and technologies. What was adopted and how delivery systems changed in response was complex, but these were crucial factors governing what applications could go forward.

Meanwhile, the number of molecular genetic tests was increasing rapidly, helping patients to make their own healthcare decisions. Fortunately, mentors and academic research training prepared Billings for this new world.

His basic science work combined with his role as a clinical provider and administrator helped him to realize that genetic screens might play an increasing role in deciding who gets what kind of treatment. "How predictive genetic technology was going to be assimilated into healthcare was the key flash point for me," he says. "It's the positive side of the

use of genetic information in healthcare."

So in 1999, he co-founded GeneSage, a genetic-services company based in San Francisco. He sees his latest appointment as vice-president and national director for genetics and genomics at the Laboratory Corporation of America in Research Triangle Park, North Carolina, as the next step, because the company is a leader in evaluating, developing and distributing genomic screening and diagnostic tests nationally.

Billings says now that his success has come about partly because he kept his interests varied — for instance, lecturing in anthropology and contributing, in the early days, to studies of the social, legal and policy implications of the Human Genome Project. That approach is contrary to what many academics now tell their students, he says. "Scientists should not feel like they are hurting themselves by having broad experiences," he explains. ■

CV

2003: Adjunct professor, anthropology dept, University of California, Berkeley

1999: Founder, GeneSage, San Francisco

1997–2000: Deputy network director and chief medical officer, VA Heart of Texas Health Care Network

1995: Founder, CBR Systems

1994–7: Associate clinical professor, Department of Medicine, Stanford University Medical School; Deputy chief of staff, VA Palo Alto Health Care System

1990–2: Vice-chairman of medicine, California Pacific Medical Center, San Francisco

1988–90: Instructor, Harvard Medical School